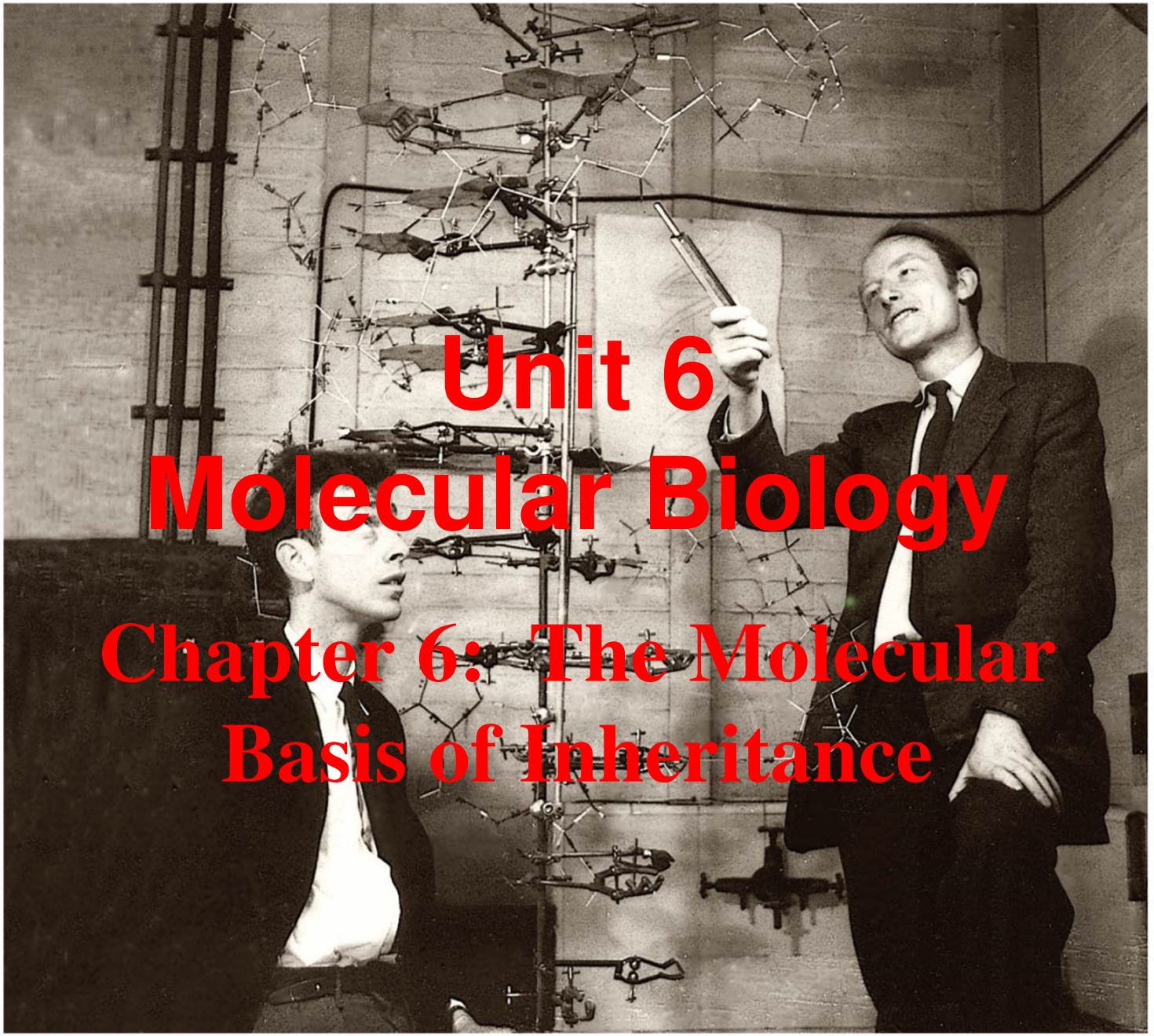


Unit 6
Molecular Biology

Chapter 6: The Molecular Basis of Inheritance

Overview: Life's Operating Instructions

- In 1953, James Watson and Francis Crick introduced an elegant double-helical model for the structure of deoxyribonucleic acid, or DNA
- Nucleic acids are unique in their ability to direct their own replication from monomers
- Hereditary information is encoded in DNA and reproduced in all cells of the body
  - DNA is copied during **DNA replication**

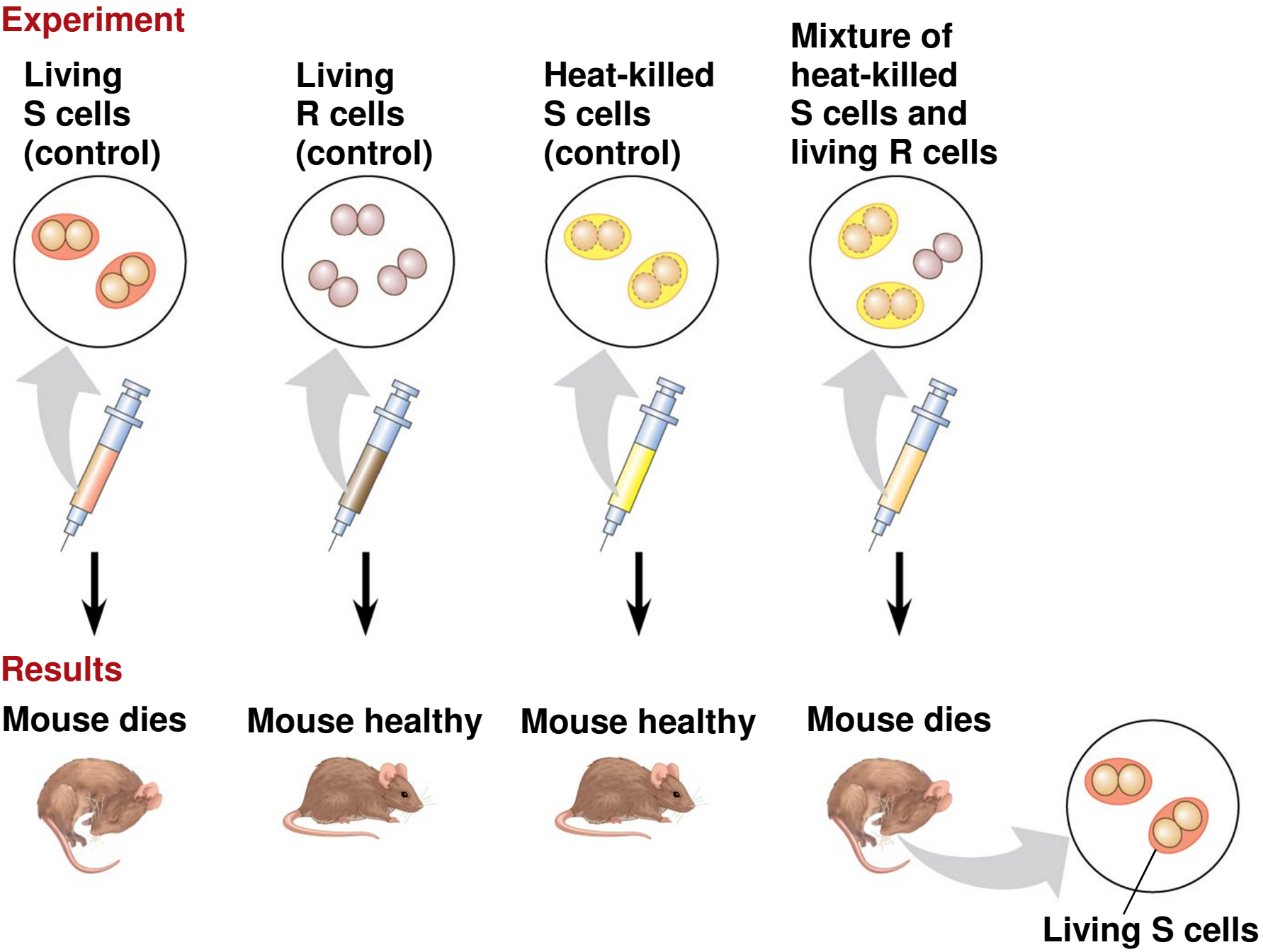
Concept 13.1: DNA is the genetic material

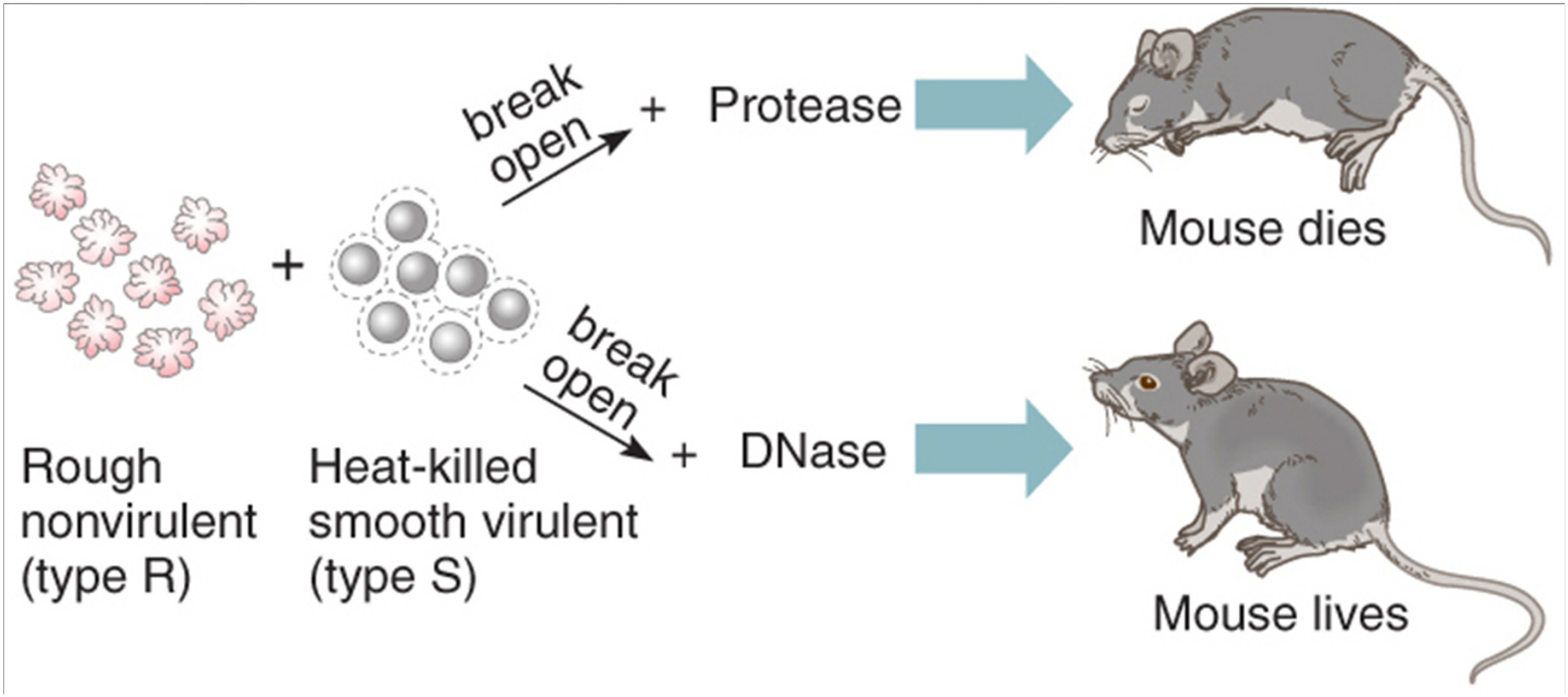
- Early in the 20th century, the identification of the molecules of inheritance loomed as a major challenge to biologists
- When T. H. Morgan's group showed that genes are located on chromosomes, the leading candidates for the genetic material became the two components of chromosomes
  - DNA
  - Protein
- The key factor in determining the genetic material was choosing appropriate experimental organisms
- The role of DNA in heredity was first discovered by studying bacteria and the viruses that infect them

Evidence That DNA Can Transform Bacteria

- In 1928, Frederick Griffith worked with two strains of a bacterium, one pathogenic and one harmless
- When he mixed heat-killed remains of the pathogenic strain with living cells of the harmless strain, some living cells became pathogenic
- He called this phenomenon **transformation**
  - Now defined as a change in genotype and phenotype due to assimilation of foreign DNA
- Later work by Oswald Avery and others identified the transforming substance as DNA

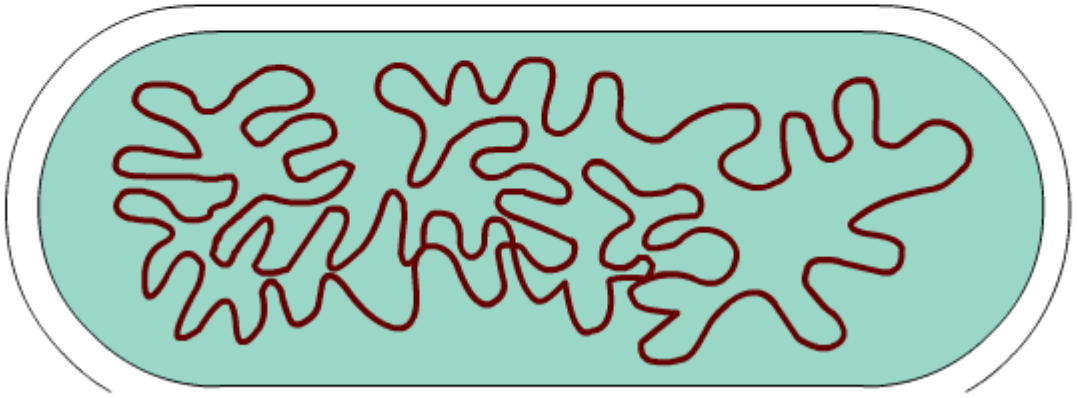
Figure 13.2


Experiment
Living S cells (control)
Living R cells (control)
Heat-killed S cells (control)
Mixture of heat-killed S cells and living R cells
Results
Mouse dies
Mouse healthy
Mouse healthy
Mouse dies
Living S cells


break open
+
Protease
Mouse dies
+
break open
+
DNase
Mouse lives
Rough nonvirulent (type R)
Heat-killed smooth virulent (type S)

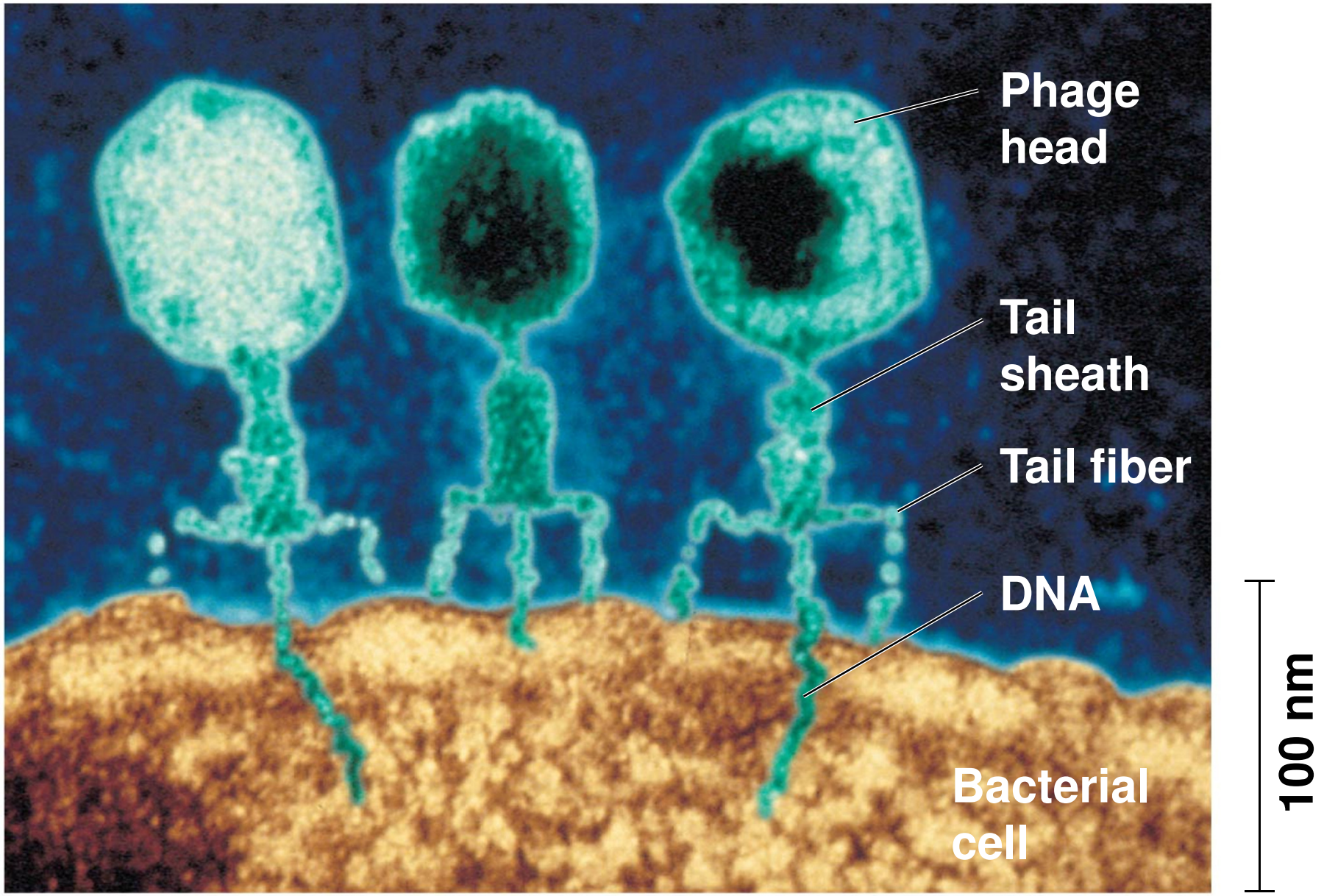
Evidence That Viral DNA Can Program Cells

- More evidence for DNA as the genetic material came from studies of **bacteriophages** (or **phages**)
  - Viruses that infect bacteria
- A virus is DNA (or RNA) enclosed by a protective protein coat
- Viruses must infect cells and take over the cells' metabolic machinery in order to reproduce



Animation: Phage T2 Reproduction
Right click slide / Select play

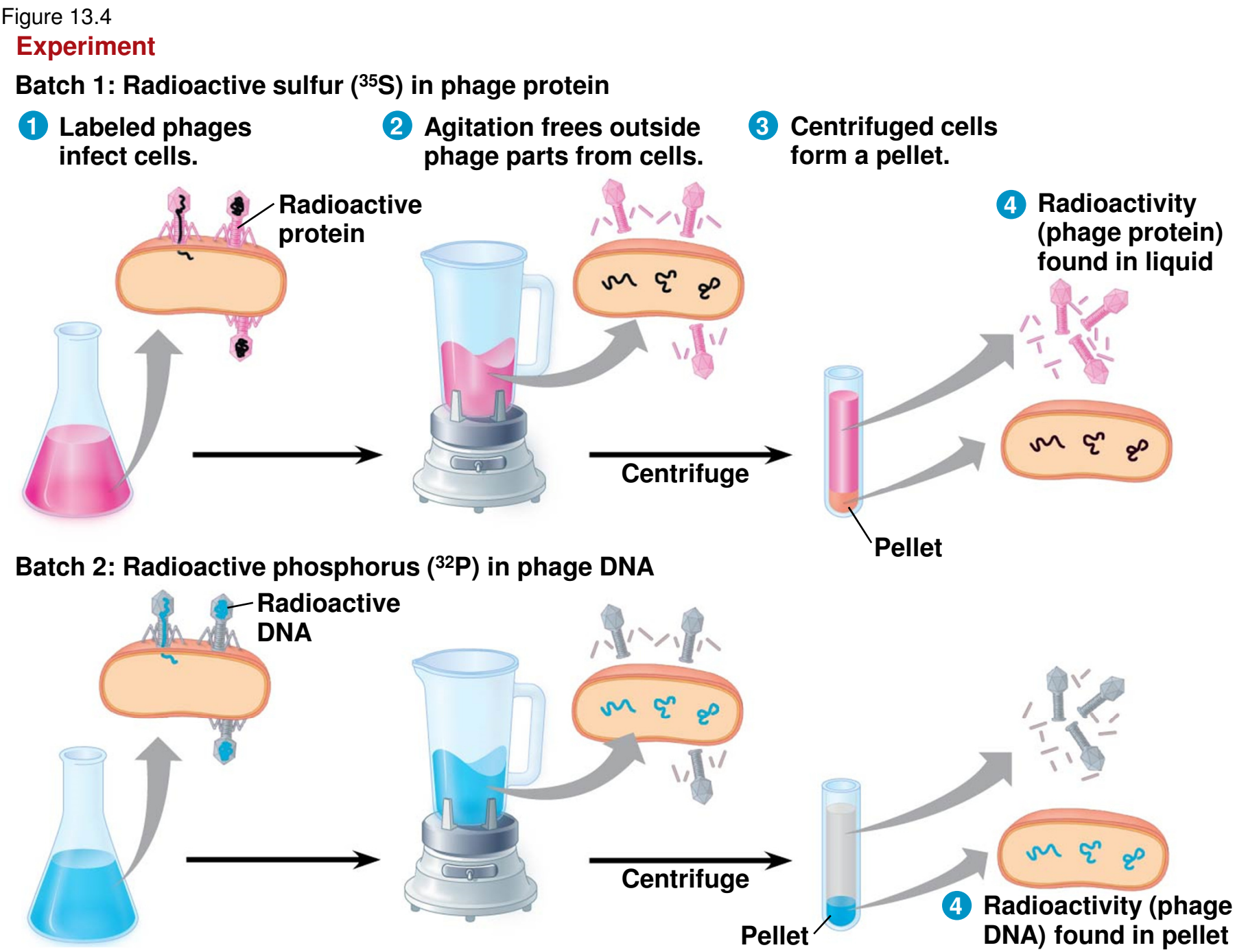
Figure 13.3



- In 1952, Alfred Hershey and Martha Chase showed that DNA is the genetic material of a phage known as T2
  - Designed an experiment showing that only the DNA of the T2 phage, and not the protein, enters an *E. coli* cell during infection
  - Used radioactive isotopes of
    - Sulfur to tag protein
    - Phosphorus to tag DNA
- They concluded that the injected DNA of the phage provides the genetic information
  - Not Protein!

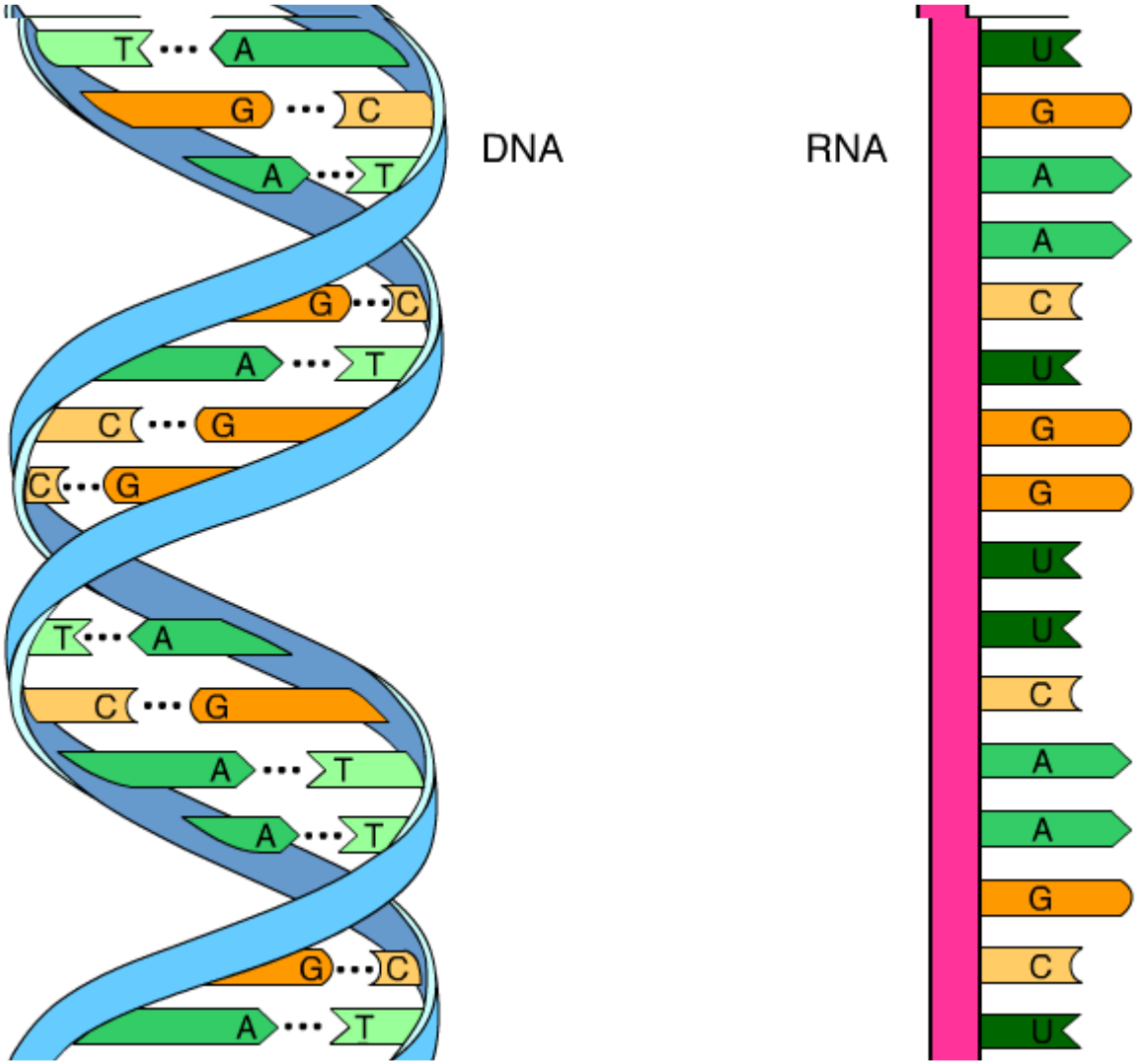
Animation: Hershey-Chase Experiment
Right click slide / Select play

Figure 13.4


Experiment
Batch 1: Radioactive sulfur (^{35}S) in phage protein
1 Labeled phages infect cells.
Radioactive protein
2 Agitation frees outside phage parts from cells.
3 Centrifuged cells form a pellet.
Centrifuge
Pellet
4 Radioactivity (phage protein) found in liquid
Batch 2: Radioactive phosphorus (^{32}P) in phage DNA
Radioactive DNA
Centrifuge
Pellet
4 Radioactivity (phage DNA) found in pellet

Additional Evidence That DNA Is the Genetic Material

- It was known that DNA is a polymer of nucleotides, each consisting of 3 components:
  - A nitrogenous base
  - A sugar
  - A phosphate group
- In 1950, Erwin Chargaff reported that DNA composition varies from one species to the next
- This evidence of diversity made DNA a more credible candidate for the genetic material



Animation: DNA and RNA Structure
Right click slide / Select play

Figure 13.5a

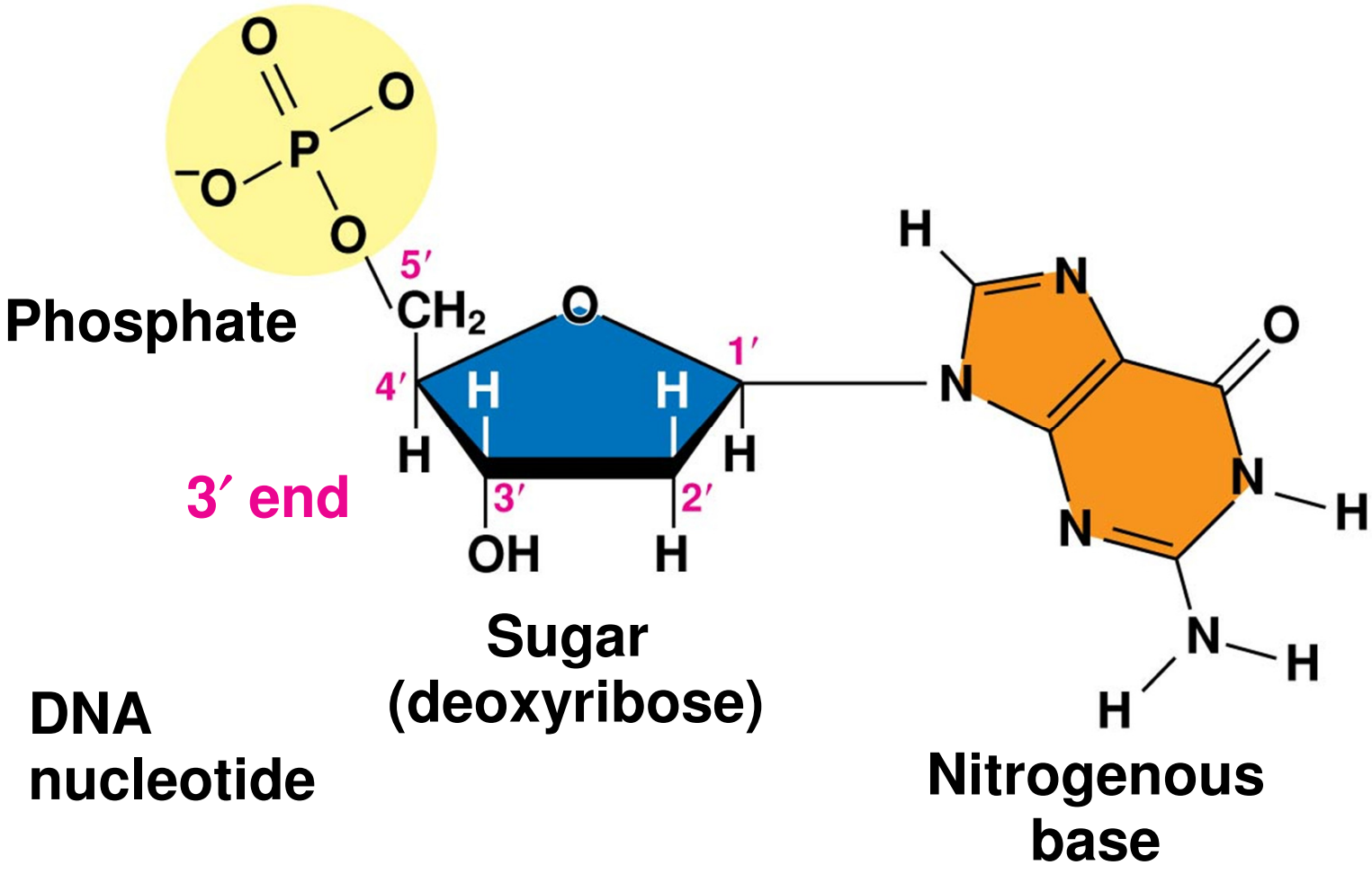
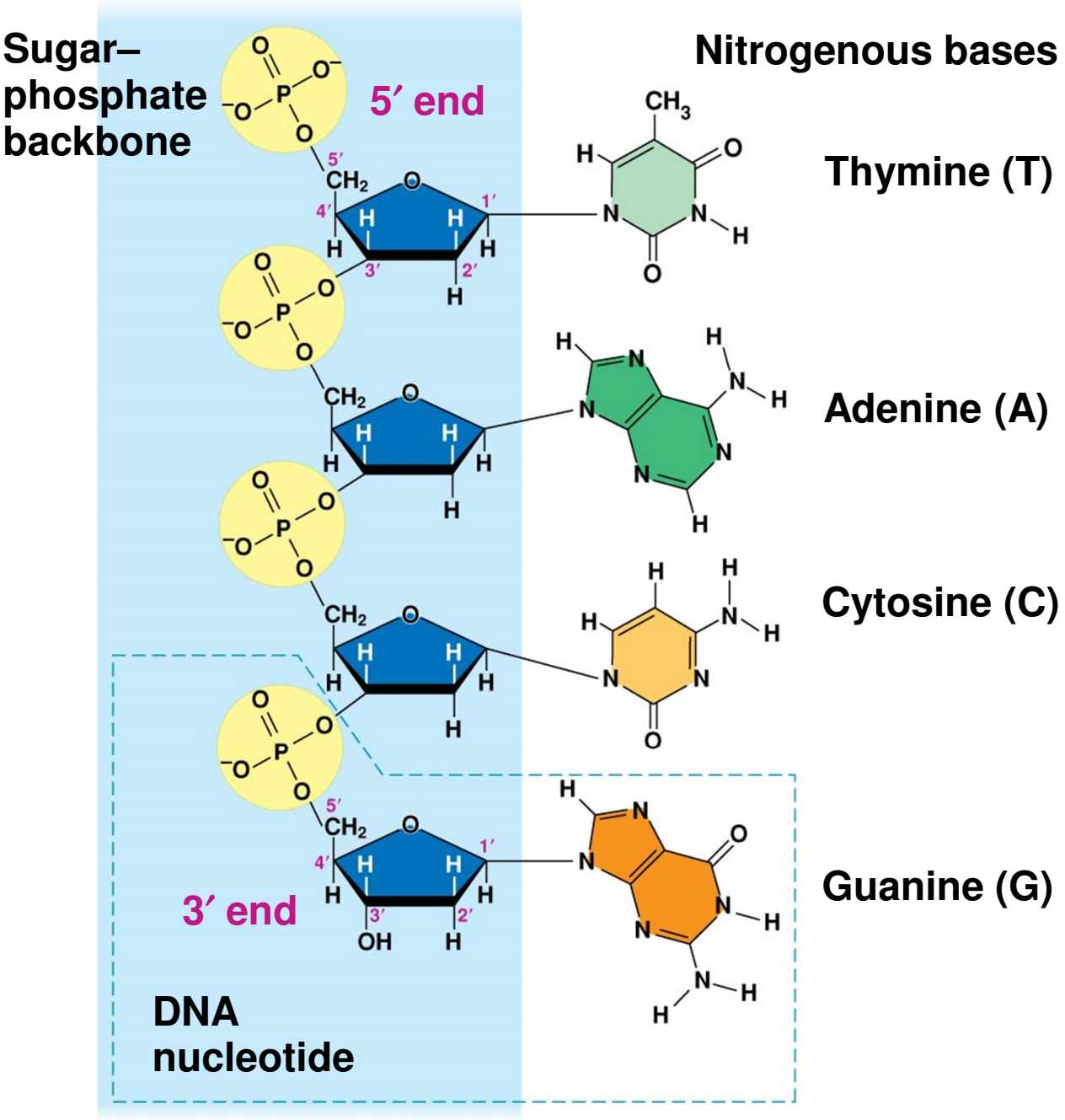

Phosphate
3′ end
DNA nucleotide
Sugar (deoxyribose)
Nitrogenous base
O
P
O
O
O
5′
CH2
4′
1′
H
H
H
H
3′
2′
OH
H
N
O
N
N
H
N
N
H
H

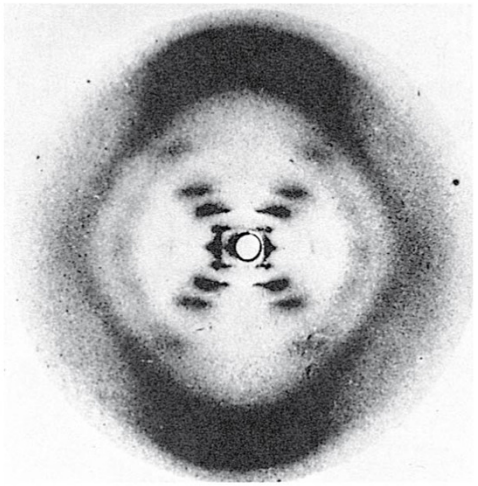
Figure 13.5


Sugar–phosphate backbone
5′ end
Nitrogenous bases
Thymine (T)
Adenine (A)
Cytosine (C)
Guanine (G)
3′ end
DNA nucleotide

- Two findings became known as Chargaff's rules
  - The base composition of DNA varies between species
  - In any species the number of A and T bases is equal and the number of G and C bases is equal
- The basis for these rules was not understood until the discovery of the double helix

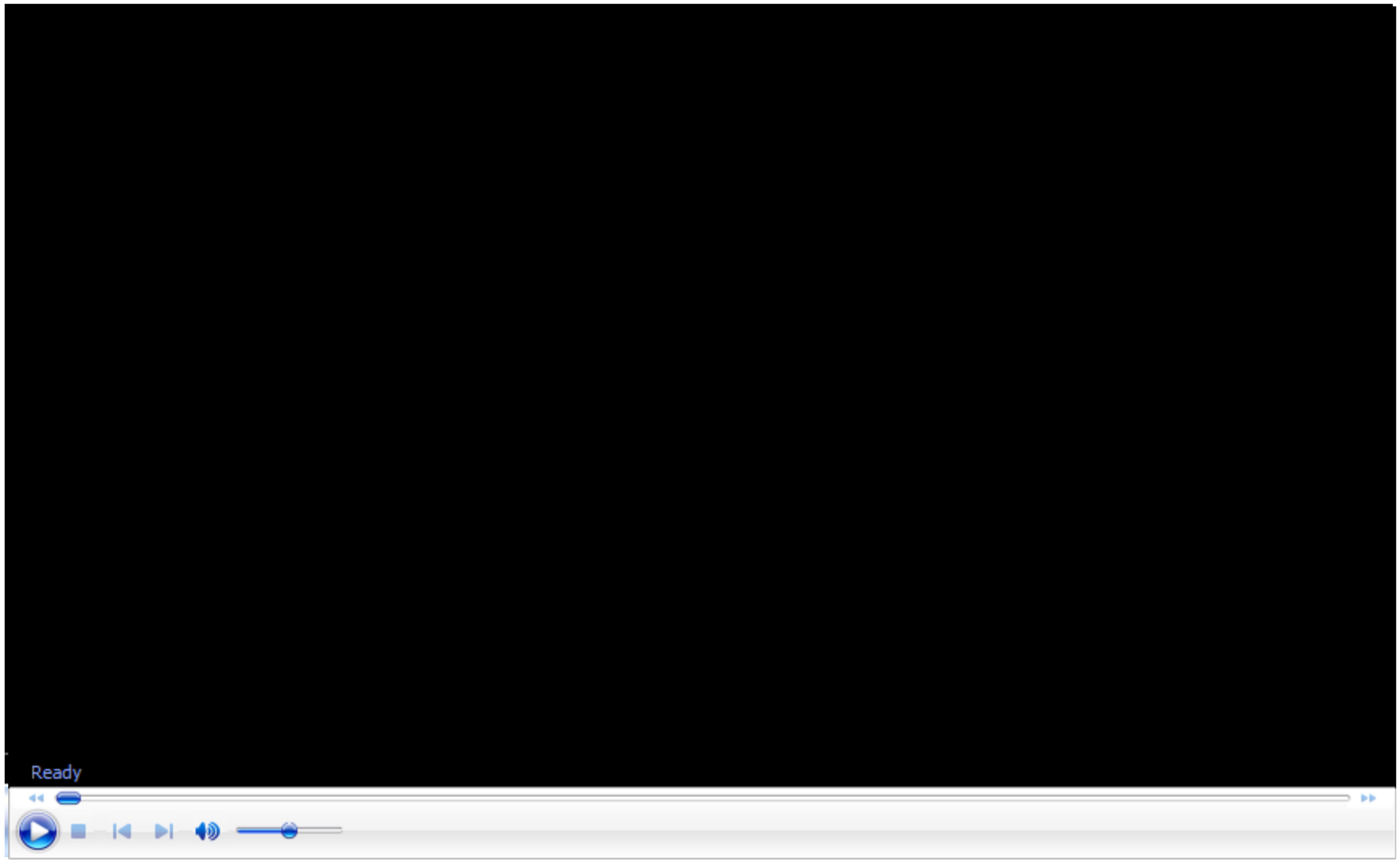
Building a Structural Model of DNA: *Scientific Inquiry*

- James Watson and Francis Crick were first to determine the structure of DNA
- Maurice Wilkins and Rosalind Franklin were using a technique called X-ray crystallography to study molecular structure
- Franklin produced a picture of the DNA molecule using this technique



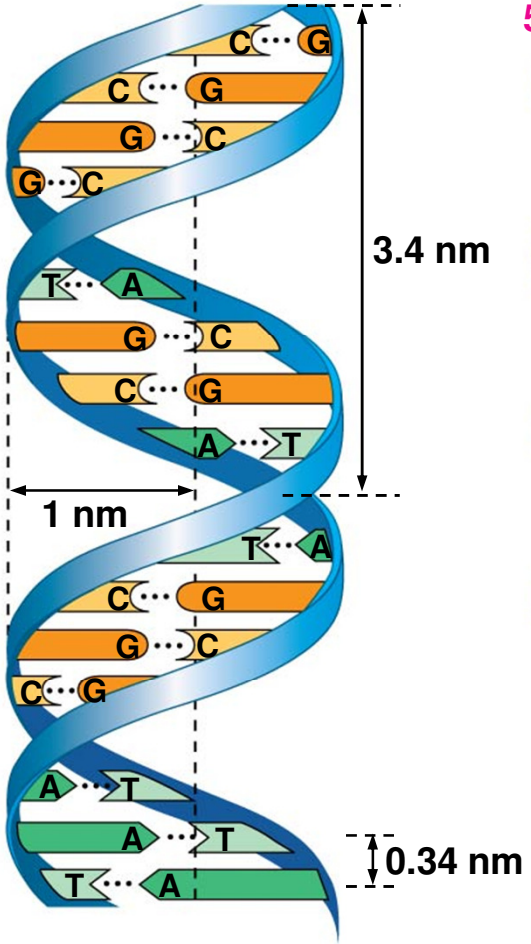
- Franklin's X-ray crystallographic images of DNA enabled Watson to deduce that DNA was helical
- The X-ray images also enabled Watson to deduce the width of the helix and the spacing of the nitrogenous bases
- The pattern in the photo suggested that the DNA molecule was made up of two strands, forming a **double helix**

Animation: DNA Double Helix
Right click slide / Select play

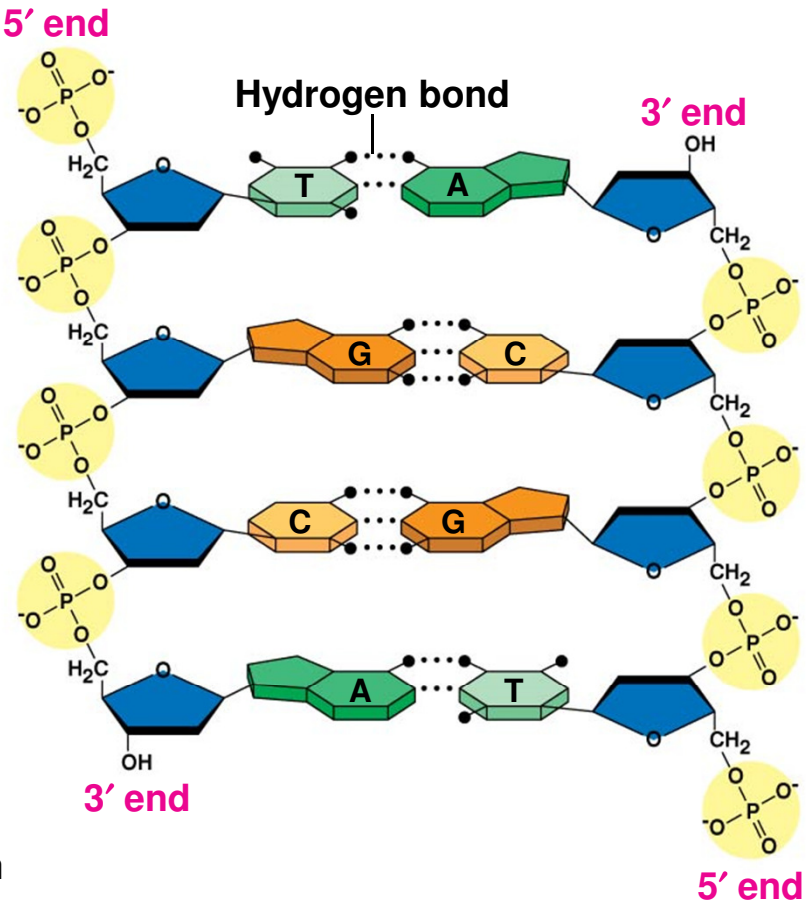


Video: DNA Surface Model

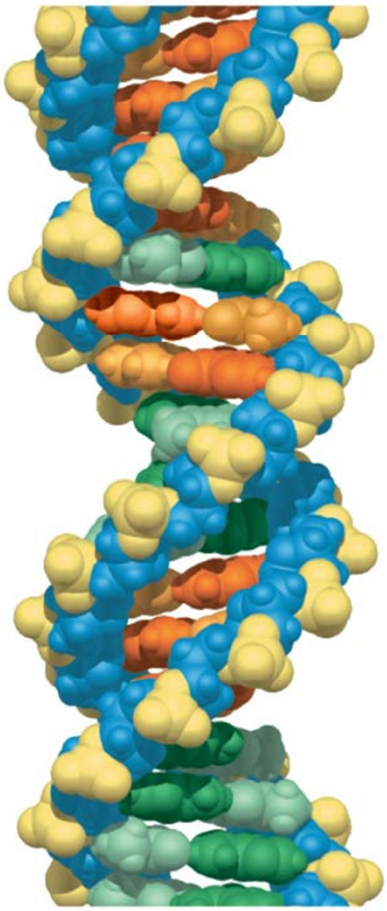
Figure 13.7



(a) Key features of DNA structure



(b) Partial chemical structure

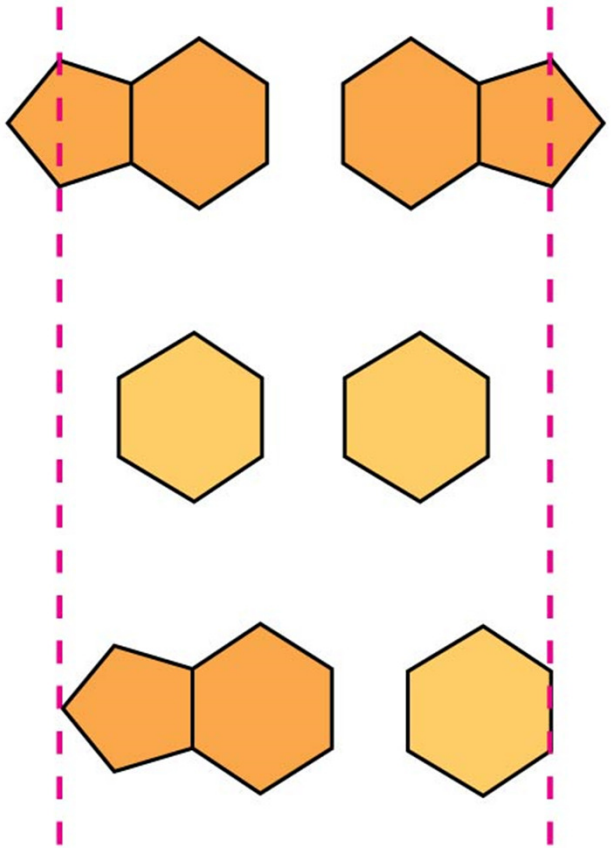


(c) Space-filling model

- Watson and Crick built models of a double helix to conform to the X-ray measurements and the chemistry of DNA
- Franklin had concluded that there were
  - Two outer sugar-phosphate backbones
  - Nitrogenous bases paired in the molecule's interior
- Watson built a model in which the backbones were **antiparallel**
  - Their subunits run in opposite directions

- At first, Watson and Crick thought the bases paired like with like (A with A, and so on)
  - But such pairings did not result in a uniform width
- Instead, pairing a *purine* with a *pyrimidine* resulted in a uniform width consistent with the X-ray data
  - Purines have 2 organic rings
    - Guanine
    - Adenine
  - Pyrimidines have single ring
    - Cytosine
    - Thymine

Figure 13.UN02



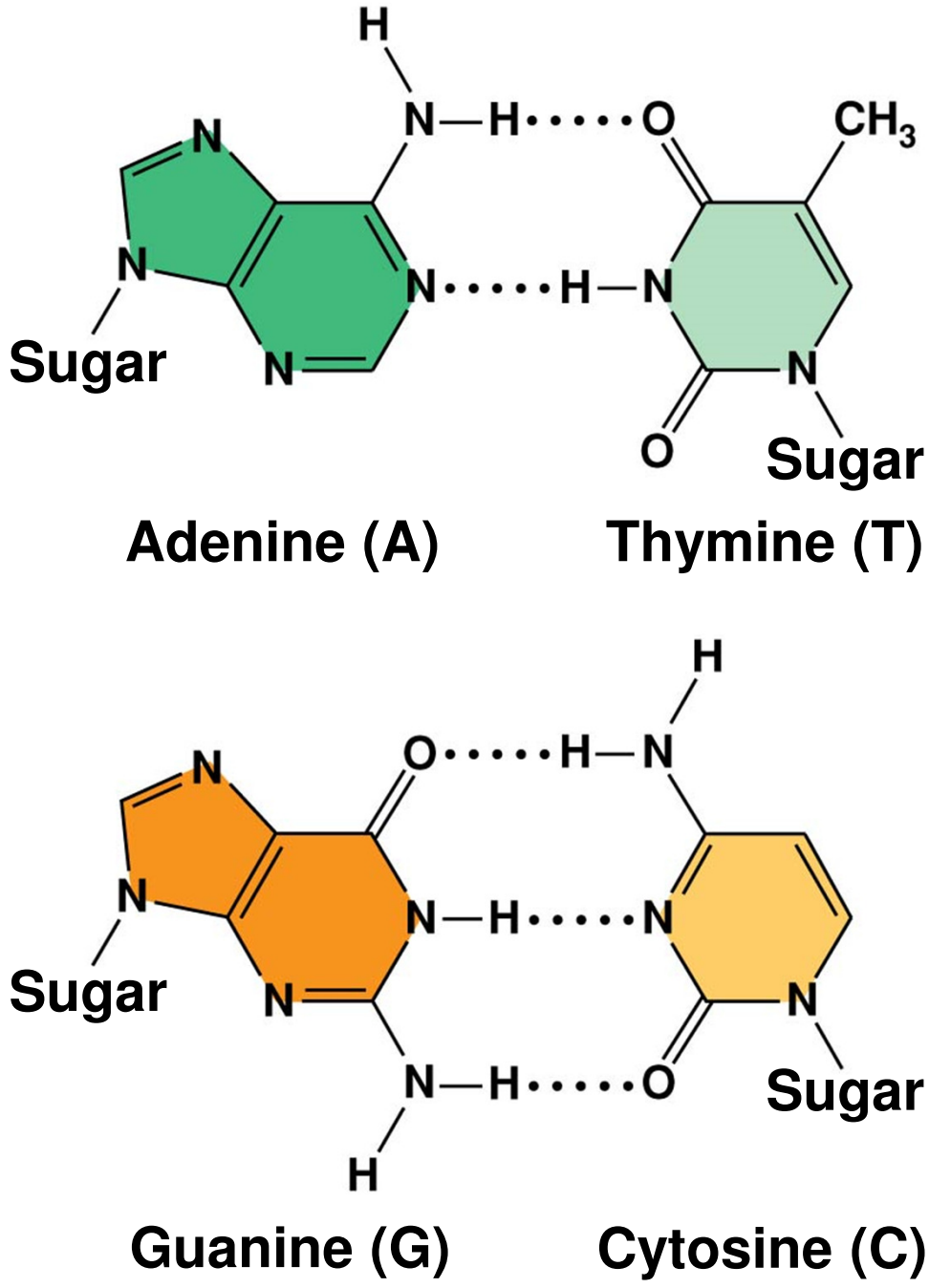
Purine + purine: too wide

Pyrimidine + pyrimidine: too narrow

Purine + pyrimidine: width consistent with X-ray data

- Watson and Crick reasoned that the pairing was more specific, dictated by the base structures
- They determined that
  - Adenine (A) paired only with thymine (T)
    - Form 2 hydrogen bonds
  - Guanine (G) paired only with cytosine (C)
    - Form 3 hydrogen bonds

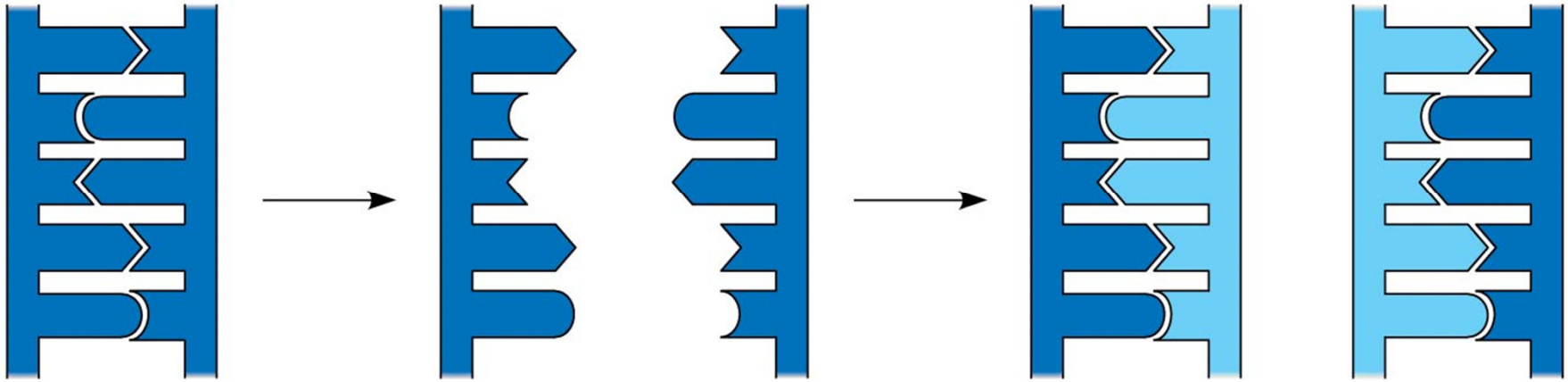
Figure 13.8


H
N—H·····O
CH3
N
N
N·····H—N
Sugar
N
N
O
Sugar
Adenine (A)
Thymine (T)
H
O·····H—N
N
N
N—H·····N
Sugar
N
N
N
N—H·····O
Sugar
H
Guanine (G)
Cytosine (C)

- The Watson-Crick model explains Chargaff's rules:
  - In any organism
    - The amount of A = T
    - The amount of G = C
- Base pairing rules do NOT restrict the sequence of nucleotides along each DNA strand
  - Linear sequence of 4 bases can be varied in countless ways
  - Each gene has a unique order, or base sequence

Concept 13.2: Many proteins work together in DNA replication and repair

- The relationship between structure and function is manifest in the double helix
- Watson and Crick noted that the specific base pairing suggested a possible copying mechanism for genetic material



(a) Parental molecule

(b) Separation of parental strands into templates

(c) Formation of new strands complementary to template strands

The Basic Principle: Base Pairing to a Template Strand

- Since the two strands of DNA are *complementary*, each strand acts as a template for building a new strand in replication
- In DNA replication, the parent molecule unwinds, and two new daughter strands are built based on base-pairing rules
- Watson and Crick's prediction was for a **semiconservative model** of replication
  - When a double helix replicates, each daughter molecule will have one old strand (derived or "conserved" from the parent molecule) and one newly made strand

- Competing models were
  - The *conservative* model
    - Two parent strands rejoin
  - The *dispersive* model
    - Each strand is a mix of old and new
- Experiments by Matthew Meselson and Franklin Stahl supported the *semiconservative* model
  - Classic example of elegant experimental design

Figure 13.10

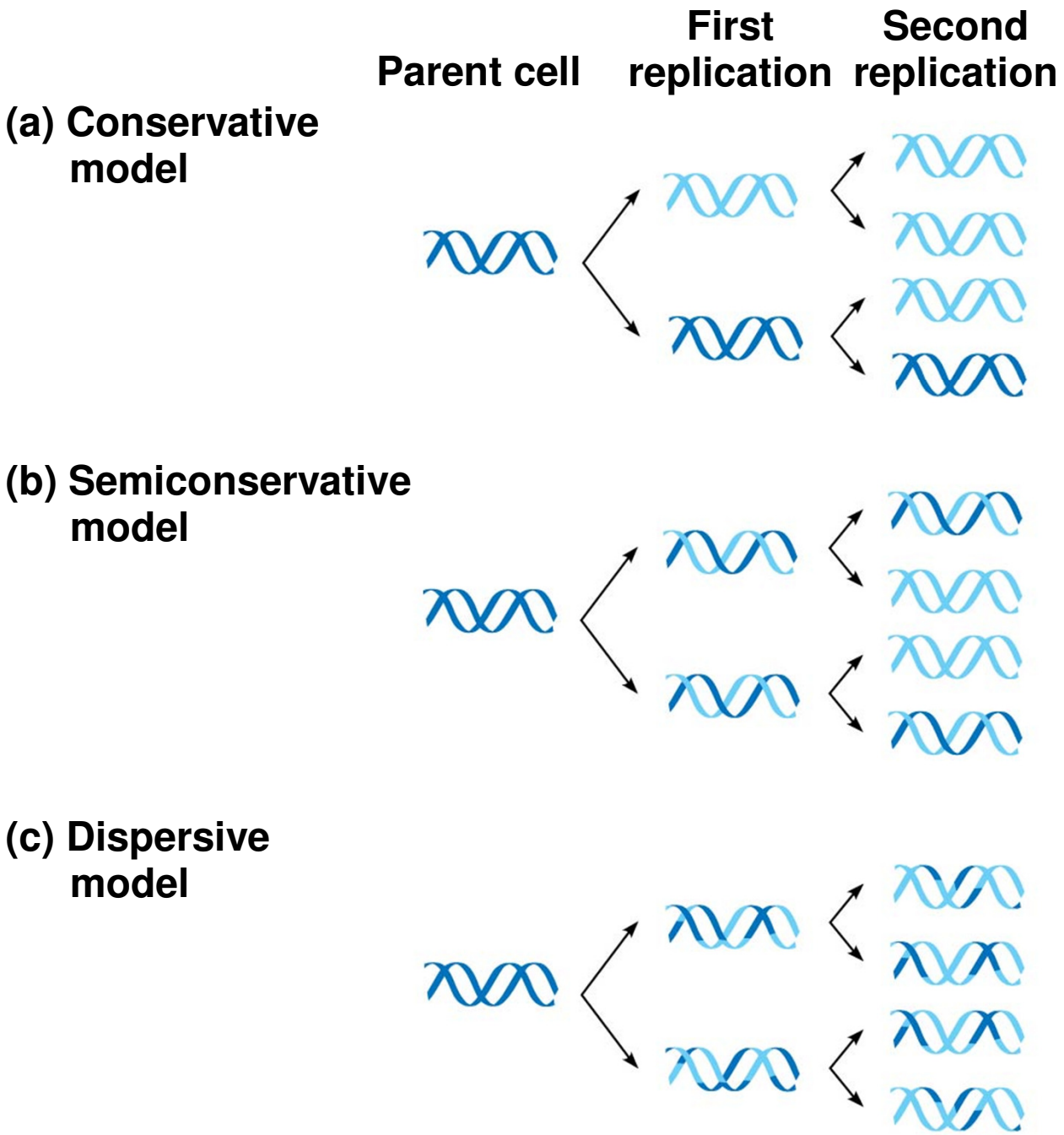

Parent cell
First replication
Second replication
(a) Conservative model
(b) Semiconservative model
(c) Dispersive model

Figure 13.11a

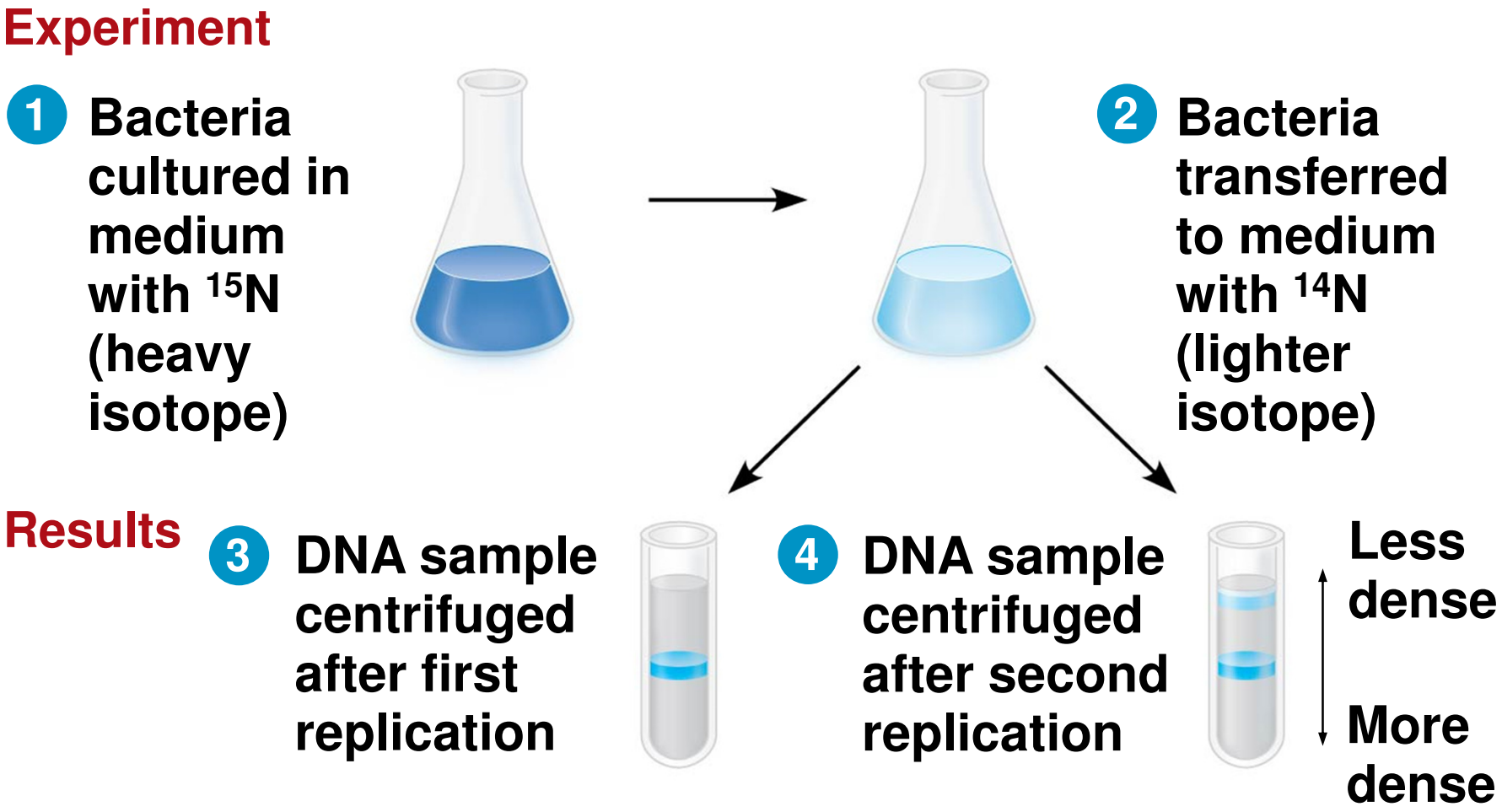
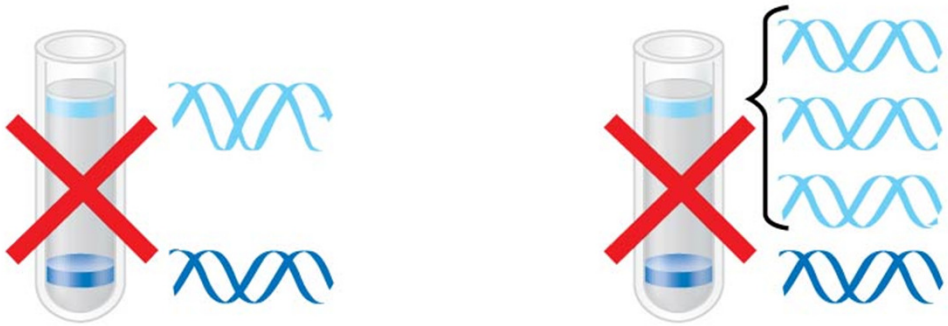


Figure 13.11b

Conclusion

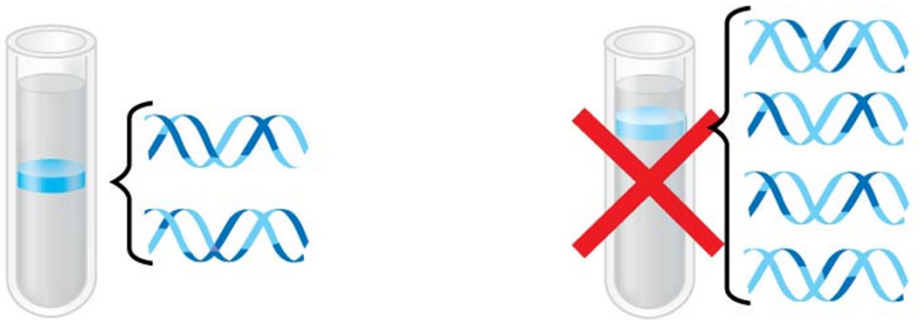
Predictions: **First replication** **Second replication**

Conservative model



Semiconservative model

Dispersive model

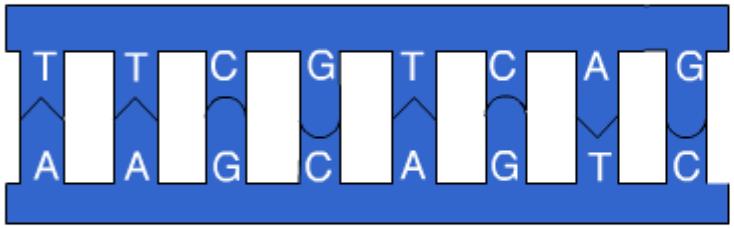


DNA Replication: *A Closer Look*

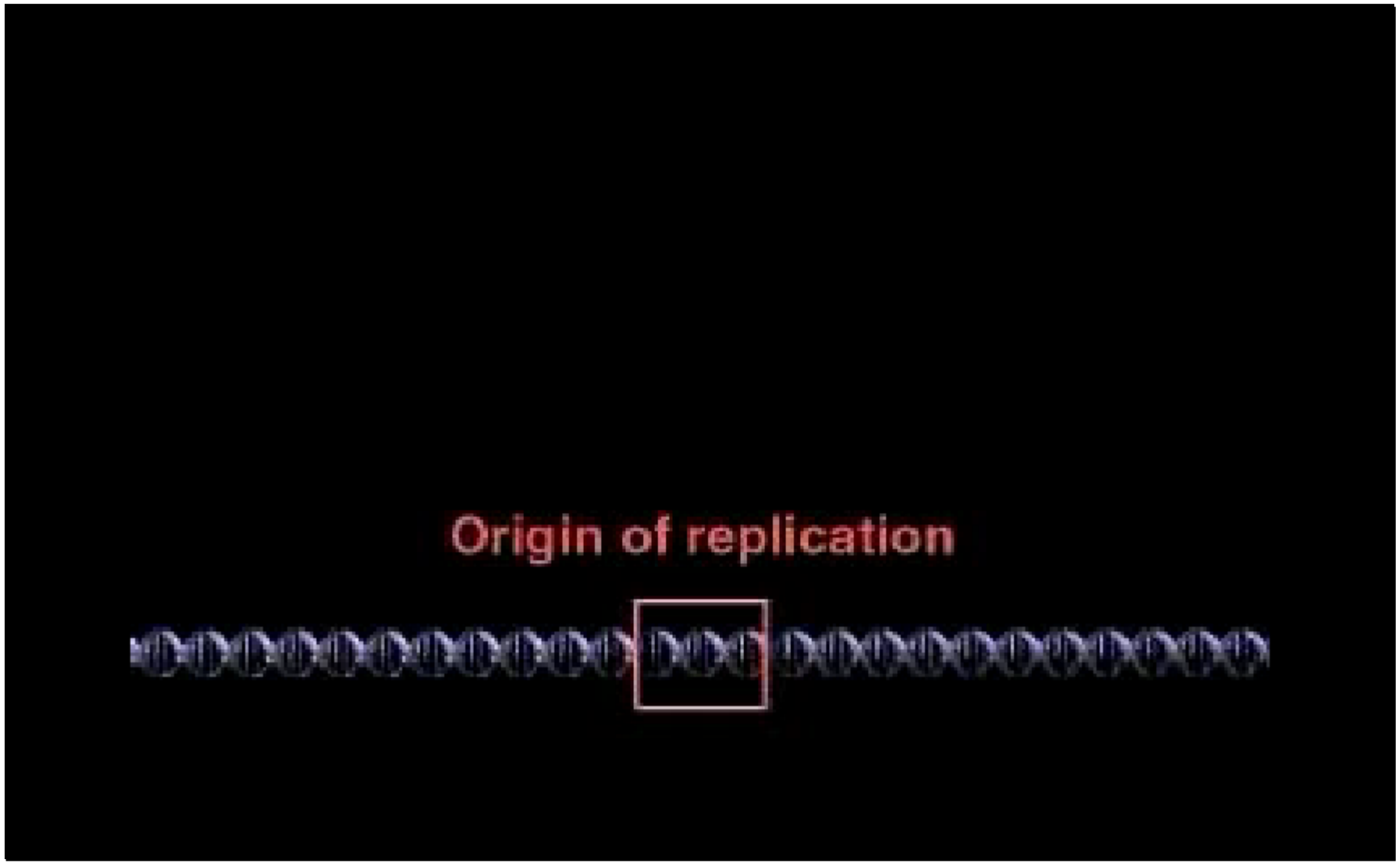
- The copying of DNA is remarkable in its speed and accuracy
- More than a dozen enzymes and other proteins participate in DNA replication
- Much more is known about how this “replication machine” works in bacteria than in eukaryotes
- Most of the process is similar between prokaryotes and eukaryotes

Getting Started

- Replication begins at particular sites called **origins of replication**
  - Here, the two DNA strands are separated, opening up a replication "bubble"
- At each end of a bubble is a **replication fork**
  - A Y-shaped region where the parental strands of DNA are being unwound

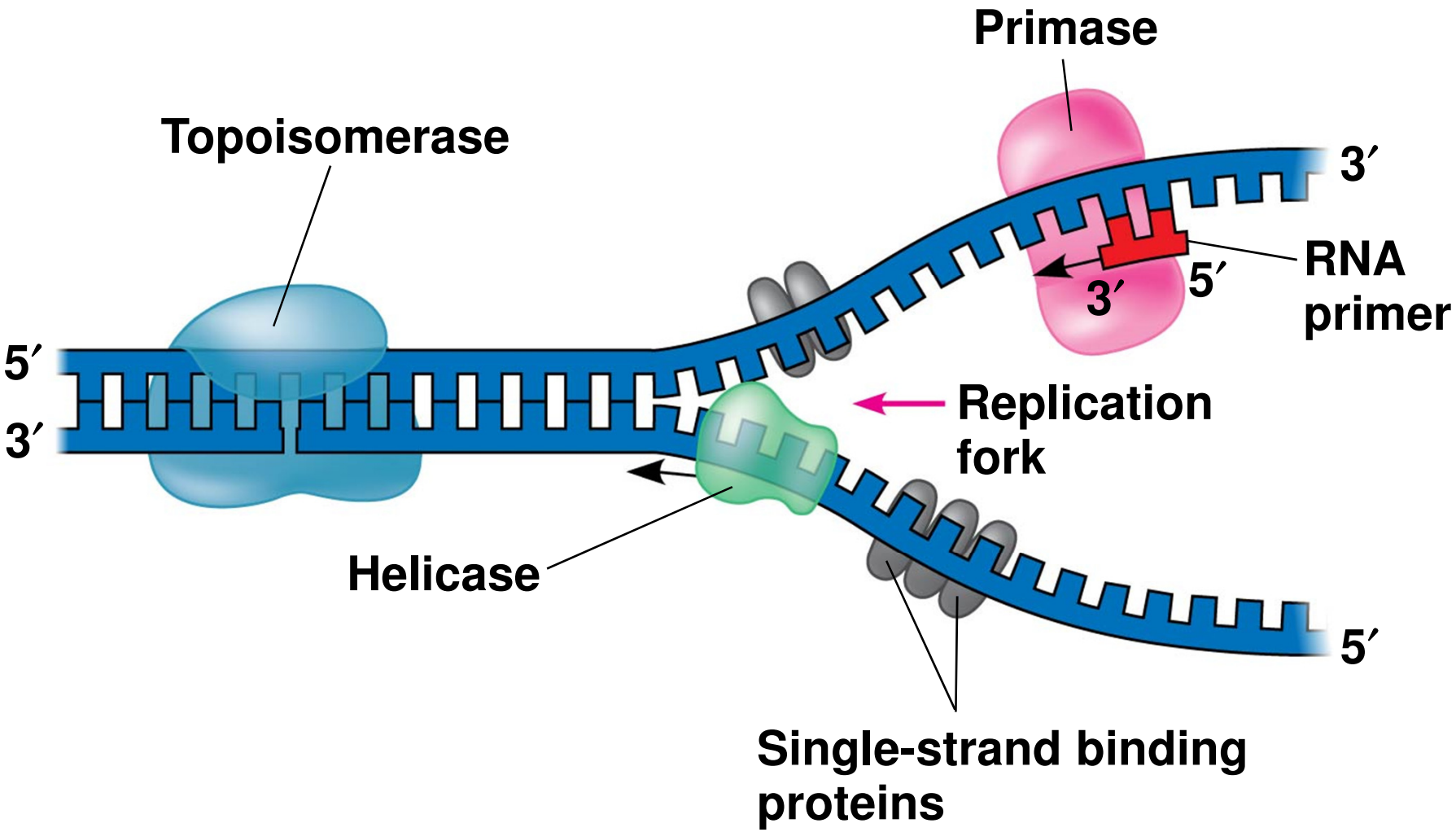


Animation: DNA Replication Overview
Right click slide / Select play



Animation: Origins of Replication
Right click slide / Select play

Figure 13.12


Primase
Topoisomerase
3′
RNA primer
5′
3′
5′
3′
Replication fork
Helicase
5′
Single-strand binding proteins

- **Helicases** are enzymes that untwist the double helix at the replication forks
- **Single-strand binding proteins** bind to and stabilize single-stranded DNA
- **Topoisomerase** relieves the strain caused by tight twisting ahead of the replication fork by breaking, swiveling, and rejoining DNA strands

- Prokaryotes
  - Typically have one circular chromosome
  - Has single origin in replication, forming one replication bubble
  - Replication proceeds in both directions until entire molecule is copied
- Eukaryotes
  - Have hundreds to thousands of replication origins
  - Multiple replication bubbles form and eventually fuse
    - Speeds up the copying of DNA
  - Replication proceeds in both directions from each origin

Figure 13.13a

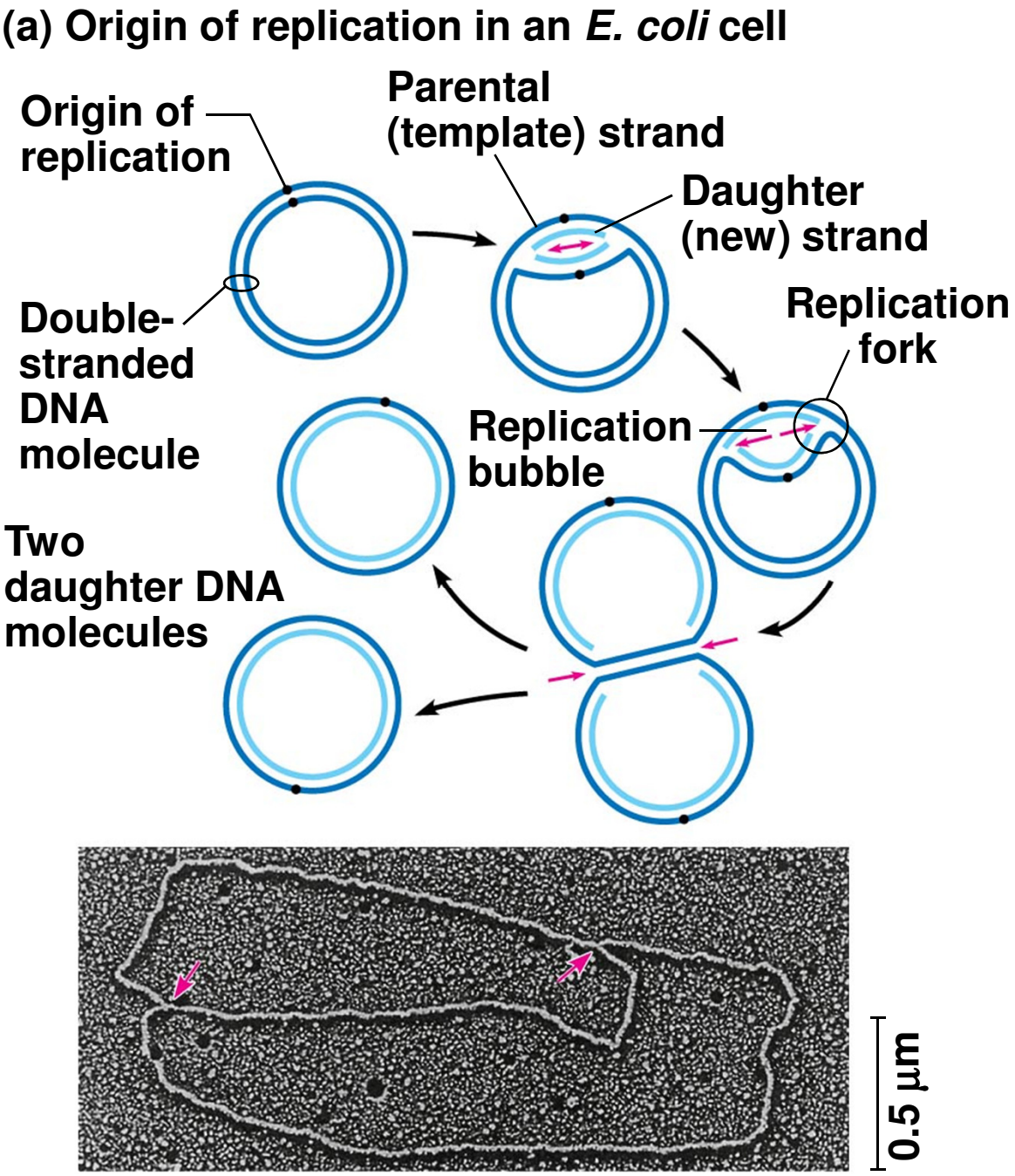
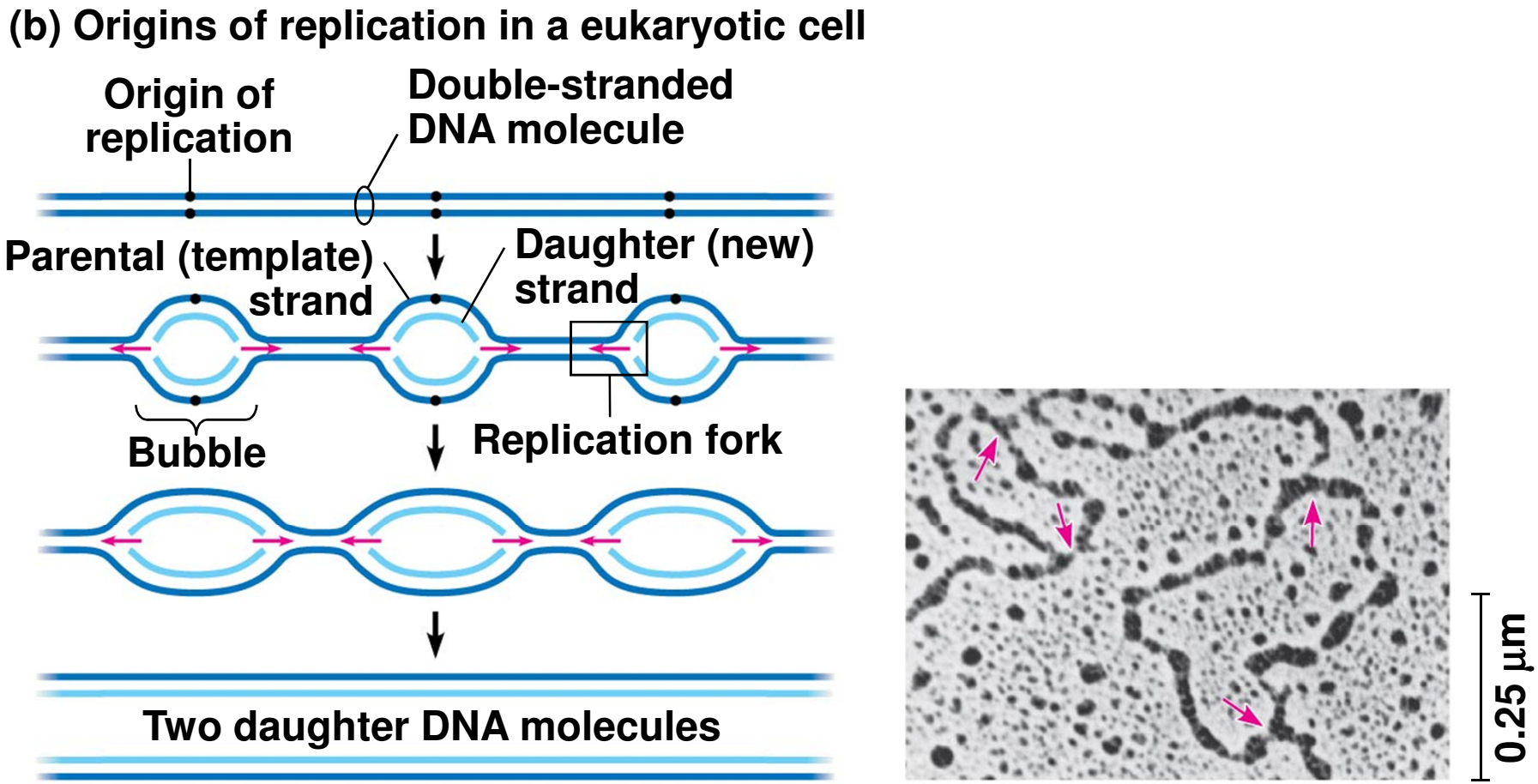


Figure 13.13b

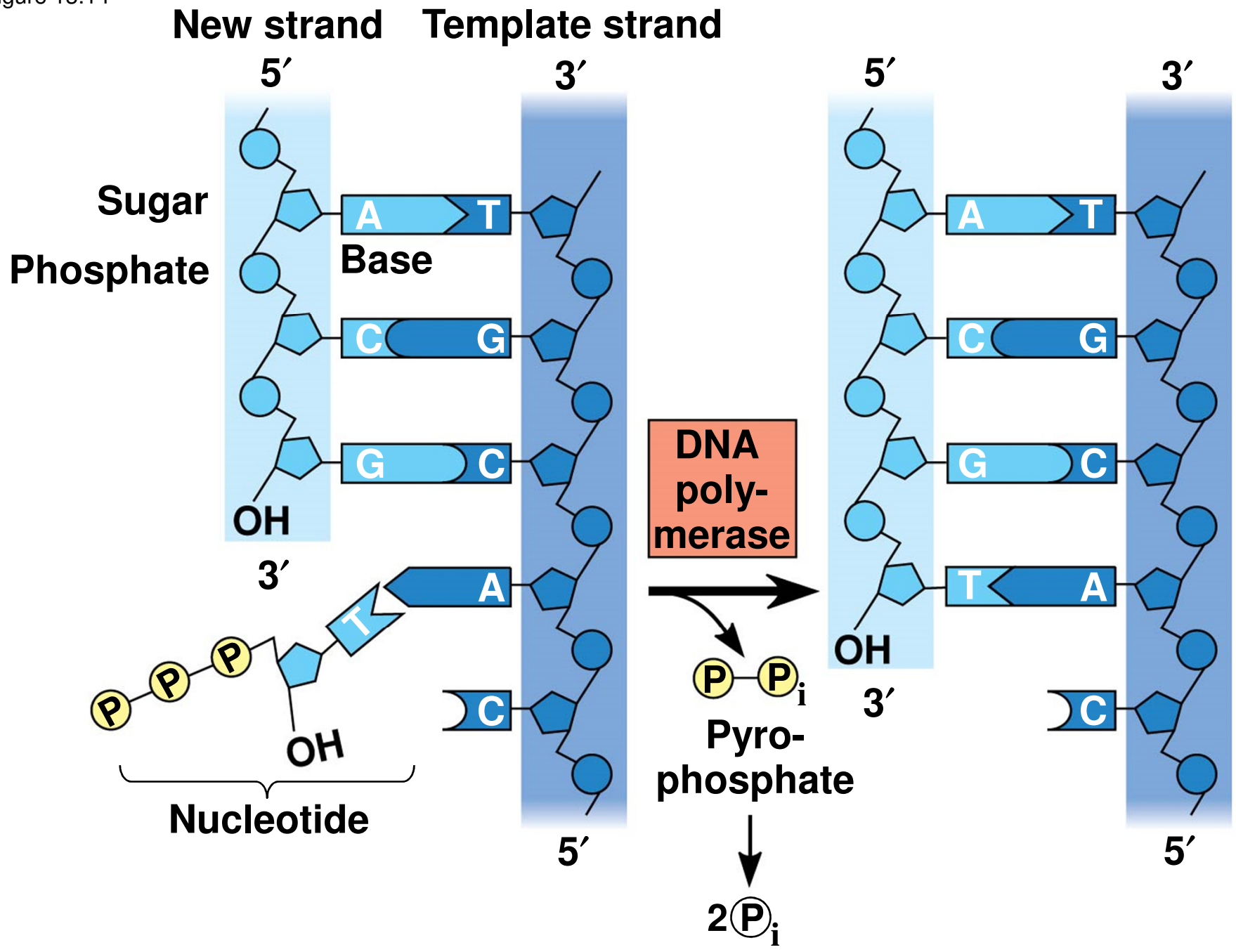


Synthesizing a New DNA Strand

- DNA polymerases cannot initiate synthesis of a polynucleotide
  - They can only add nucleotides to an already existing chain base-paired with the template
- The initial nucleotide strand is a short RNA **primer**
- The enzyme, **primase,** starts an RNA chain from a single RNA nucleotide and adds RNA nucleotides one at a time using the parental DNA as a template
- The new DNA strand will start from the 3′ end of the RNA primer

- Enzymes called **DNA polymerases** catalyze the elongation of new DNA at a replication fork
- Each nucleotide that is added to a growing DNA consists of a sugar attached to a base and three phosphate groups
- dATP is used to make DNA and is similar to the ATP of energy metabolism
  - The difference is in the sugars
    - dATP has deoxyribose
    - ATP has ribose
- As each monomer nucleotide joins the DNA strand, it loses two phosphate groups as a molecule of pyrophosphate

Figure 13.14


New strand
Template strand
5′
3′
5′
3′
Sugar
Phosphate
A
T
Base
C
G
G
C
OH
3′
A
T
P
P
P
OH
C
Nucleotide
5′
DNA poly-merase
P
P
i
Pyro-phosphate
2
P
i
5′
3′
A
T
C
G
G
C
T
A
OH
3′
C
5′

Antiparallel Elongation

- The antiparallel structure of the double helix affects replication
- DNA polymerases add nucleotides only to the free 3′ end of a growing strand
  - Therefore, a new DNA strand can elongate only in the 5′ to 3′ direction
- Along one template strand of DNA, the DNA polymerase synthesizes a **leading strand** continuously, moving toward the replication fork
- To elongate the other new strand, called the **lagging strand**, DNA polymerase must work in the direction away from the replication fork

- The lagging strand is synthesized as a series of segments called **Okazaki fragments**
- After formation of Okazaki fragments, DNA polymerase I removes the RNA primers and replaces the nucleotides with DNA
- The remaining gaps are joined together by **DNA ligase**

Animation: Leading Strand
Right click slide / Select play

Animation: Lagging Strand
Right click slide / Select play

Animation: DNA Replication Review
Right click slide / Select play

Figure 13.15a

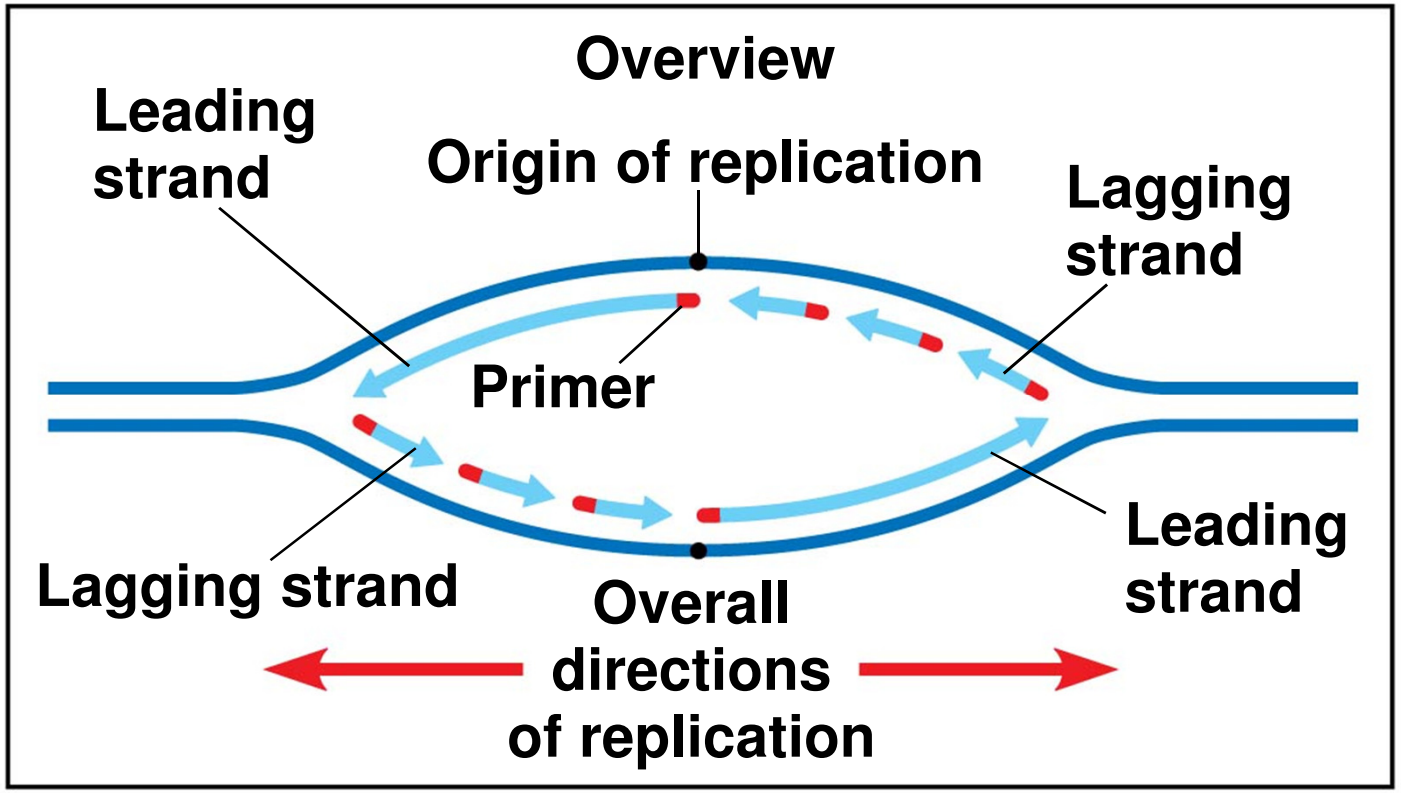


Figure 13.15b

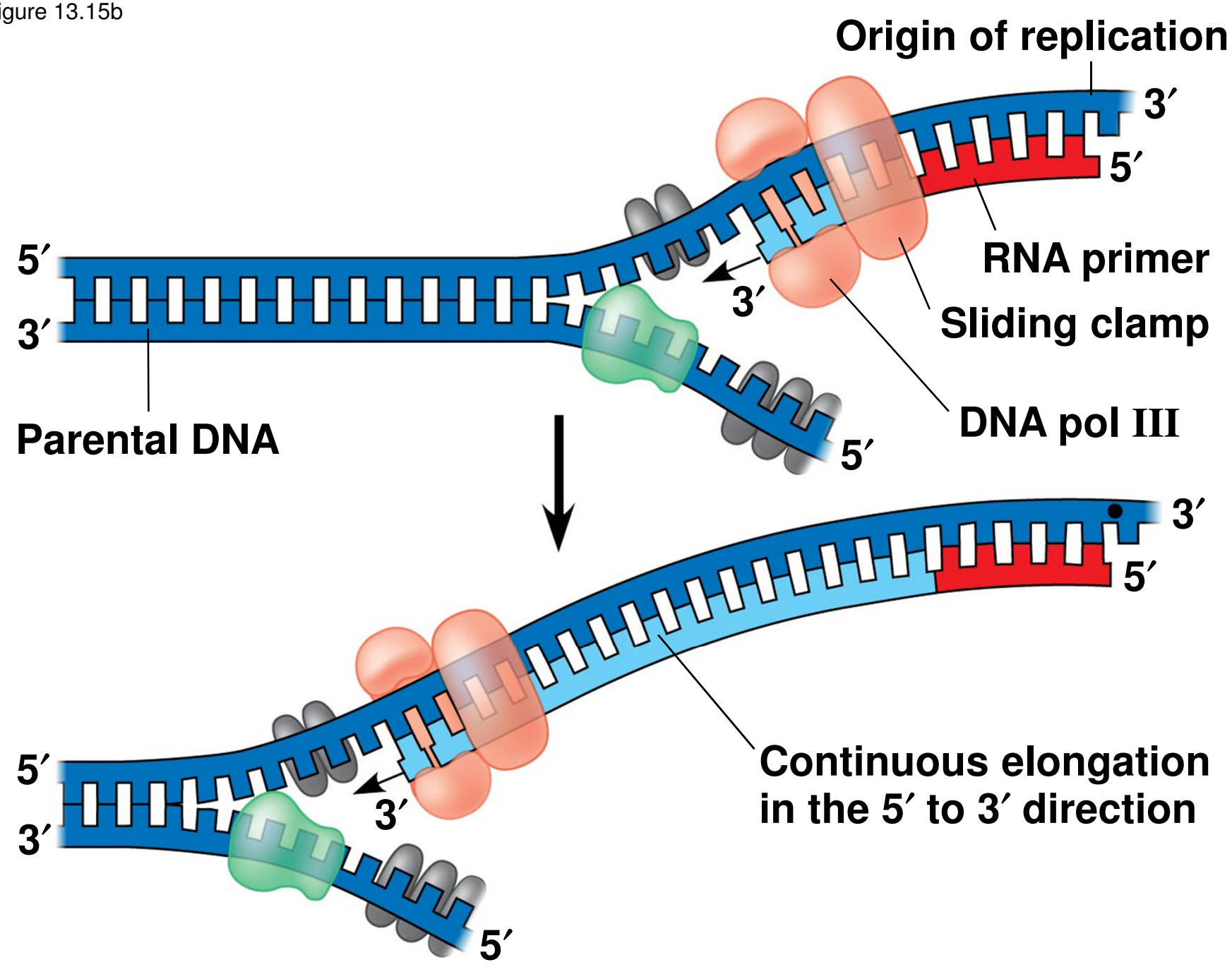


Figure 13.16b-3

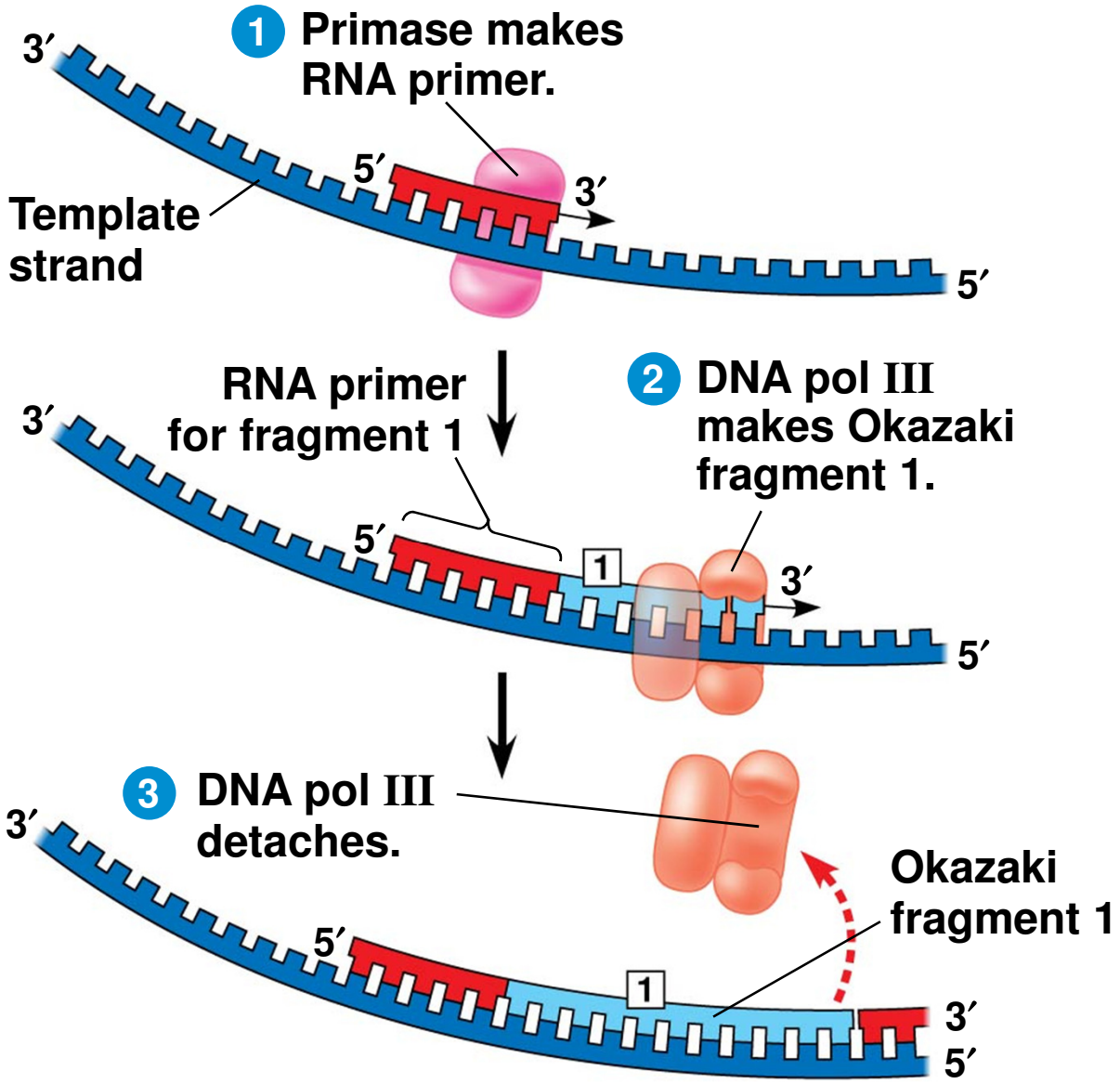


Figure 13.16c-3

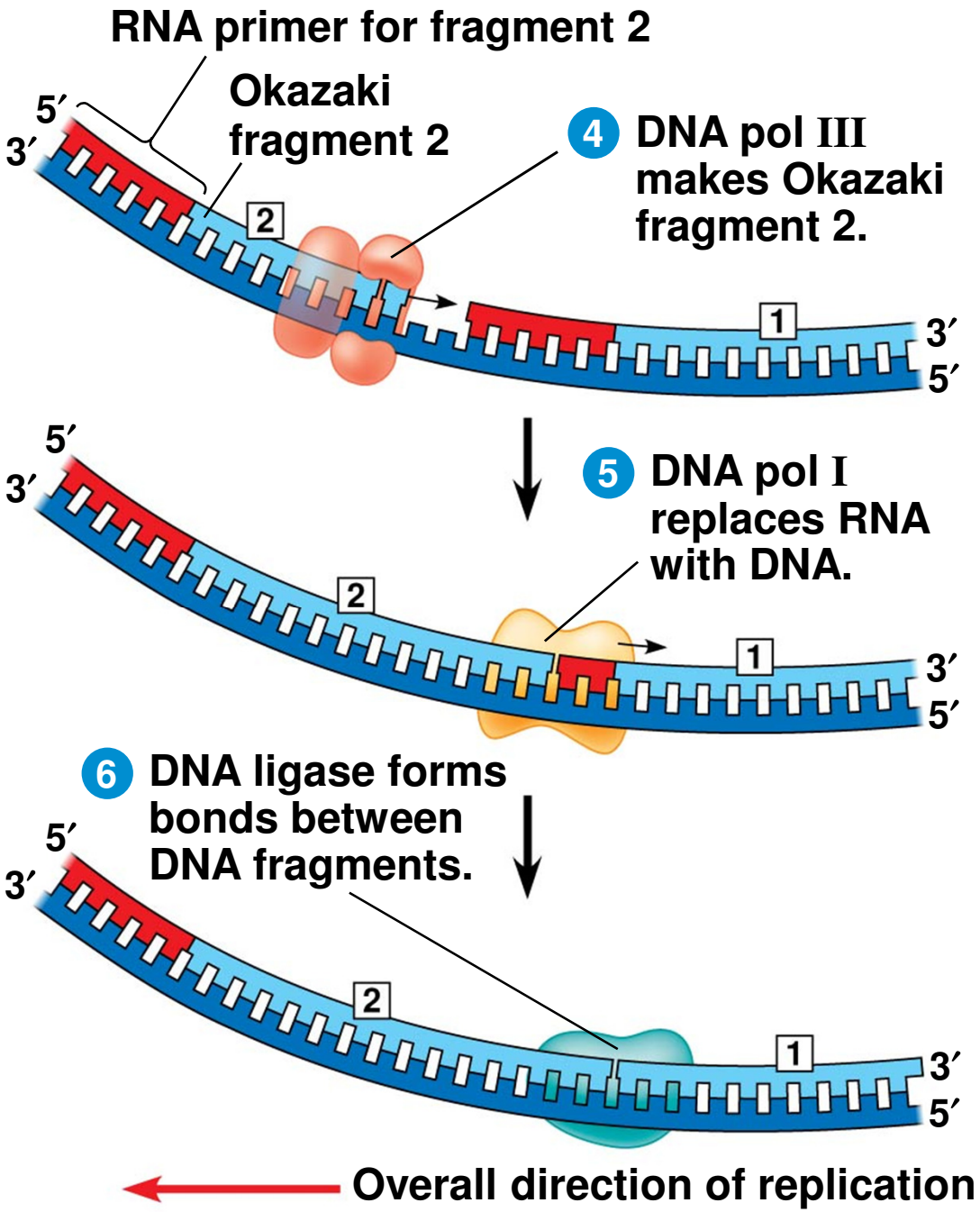
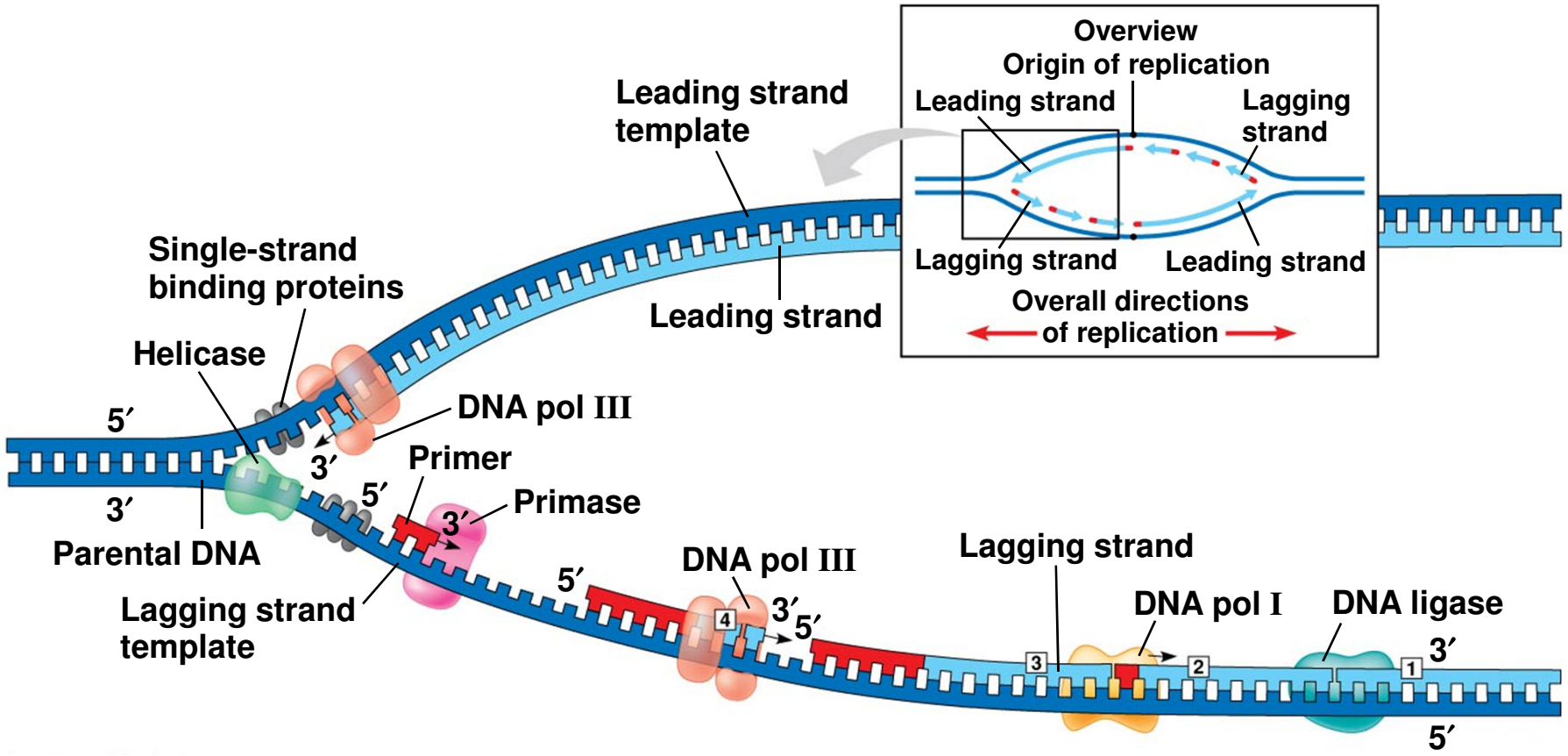


Figure 13.17

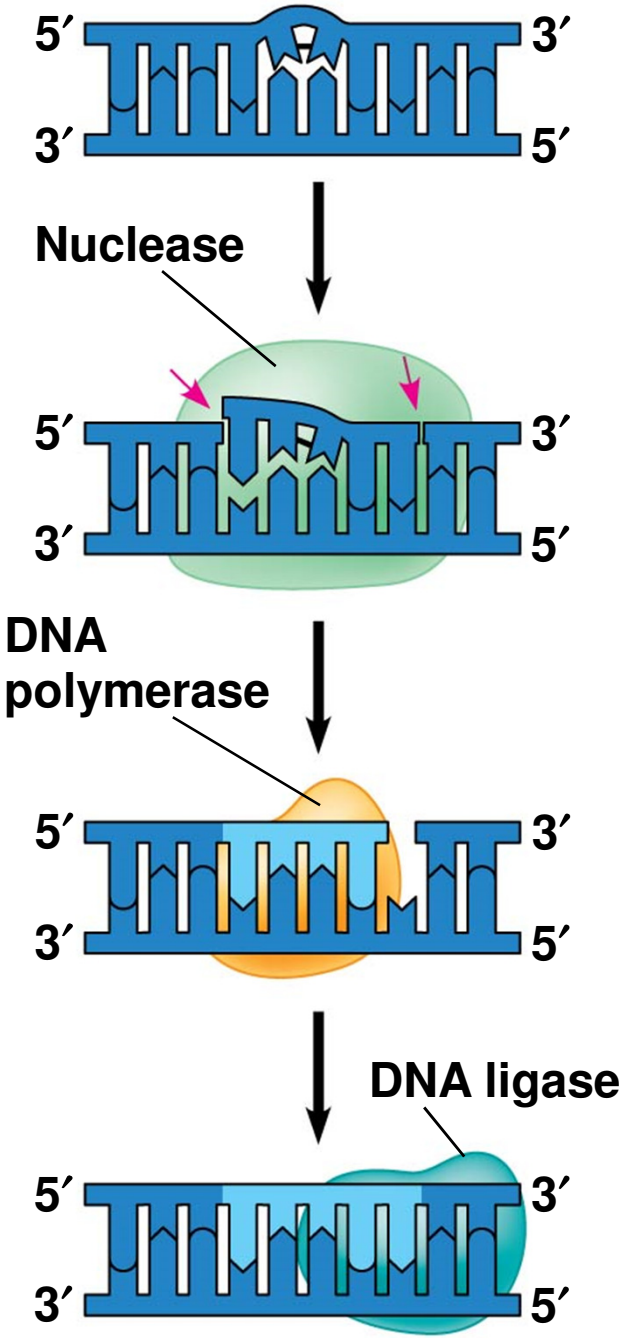

Overview
Origin of replication
Leading strand
Lagging strand
Lagging strand
Leading strand
Overall directions of replication
Leading strand template
Leading strand
Single-strand binding proteins
Helicase
DNA pol III
5′
3′
Primer
Primase
3′
5′
3′
Parental DNA
Lagging strand template
5′
DNA pol III
4
3′
5′
Lagging strand
DNA pol I
DNA ligase
3
2
1
3′
5′

Proofreading and Repairing DNA

- DNA polymerases proofread newly made DNA, replacing any incorrect nucleotides
- In **mismatch repair** of DNA, other enzymes correct errors in base pairing
  - Defect in such enzymes allows cancer-causing errors to accumulate in DNA

- DNA can be damaged by exposure to harmful chemical or physical agents such as cigarette smoke and X-rays
- DNA can also undergo spontaneous changes
  - Usually corrected before they become permanent changes (*mutations*)
- In **nucleotide excision repair**, a **nuclease** cuts out and replaces damaged stretches of DNA
  - Repair genetic damage caused by UV rays from sun
    - Ex: Thymine dimers cause DNA to buckle and interfere with DNA replication

Figure 13.19-3



Evolutionary Significance of Altered DNA Nucleotides

- Error rate after proofreading repair is low but not zero
- Sequence changes may become permanent and can be passed on to the next generation
- These changes (mutations) are the source of the genetic variation upon which natural selection operates

Replicating the Ends of DNA Molecules

- Limitations of DNA polymerase create problems for the linear DNA of eukaryotic chromosomes
- The usual replication machinery cannot complete the 5′ ends of daughter strands
- Repeated rounds of replication produce shorter DNA molecules with uneven ends

- Eukaryotic chromosomal DNA molecules have special nucleotide sequences at their ends called *telomeres*
  - Do not contain genes
  - DNA consists of multiple repetitions of one short nucleotide sequence
  - Act as a buffer zone that protects the organism's genes
- Telomeres do not prevent the shortening of DNA molecules, but they do postpone it
- It has been proposed that the shortening of telomeres is connected to aging

- If chromosomes of germ cells became shorter in every cell cycle, essential genes would eventually be missing from the gametes they produce
- An enzyme called **telomerase** catalyzes the lengthening of telomeres in germ cells
  - Not active in most human somatic cells
  - Shows inappropriate activity in some cancer cells
  - Currently under study as a target for cancer therapies

Concept 13.3: A chromosome consists of a DNA molecule packed together with proteins

- The bacterial chromosome is a double-stranded, circular DNA molecule associated with a small amount of protein
- Eukaryotic chromosomes have linear DNA molecules associated with a large amount of protein
- In a bacterium, the DNA is "supercoiled" and found in a region of the cell called the **nucleoid**
  - NOT surrounded by a membrane!

- **Chromatin** is found in the nucleus of eukaryotic cells
  - A complex of DNA and protein
- Chromosomes fit into the nucleus through an elaborate, multilevel system of packing
- As a cell prepares for mitosis, its chromatin coils and folds up (condenses)

- Chromatin packing
  - **Histone**
    - Protein around which DNA coils
  - **Nucleosome**
    - Bead-like unit of DNA packing
    - Consists of a segment of DNA wound twice around histones

Figure 13.21a

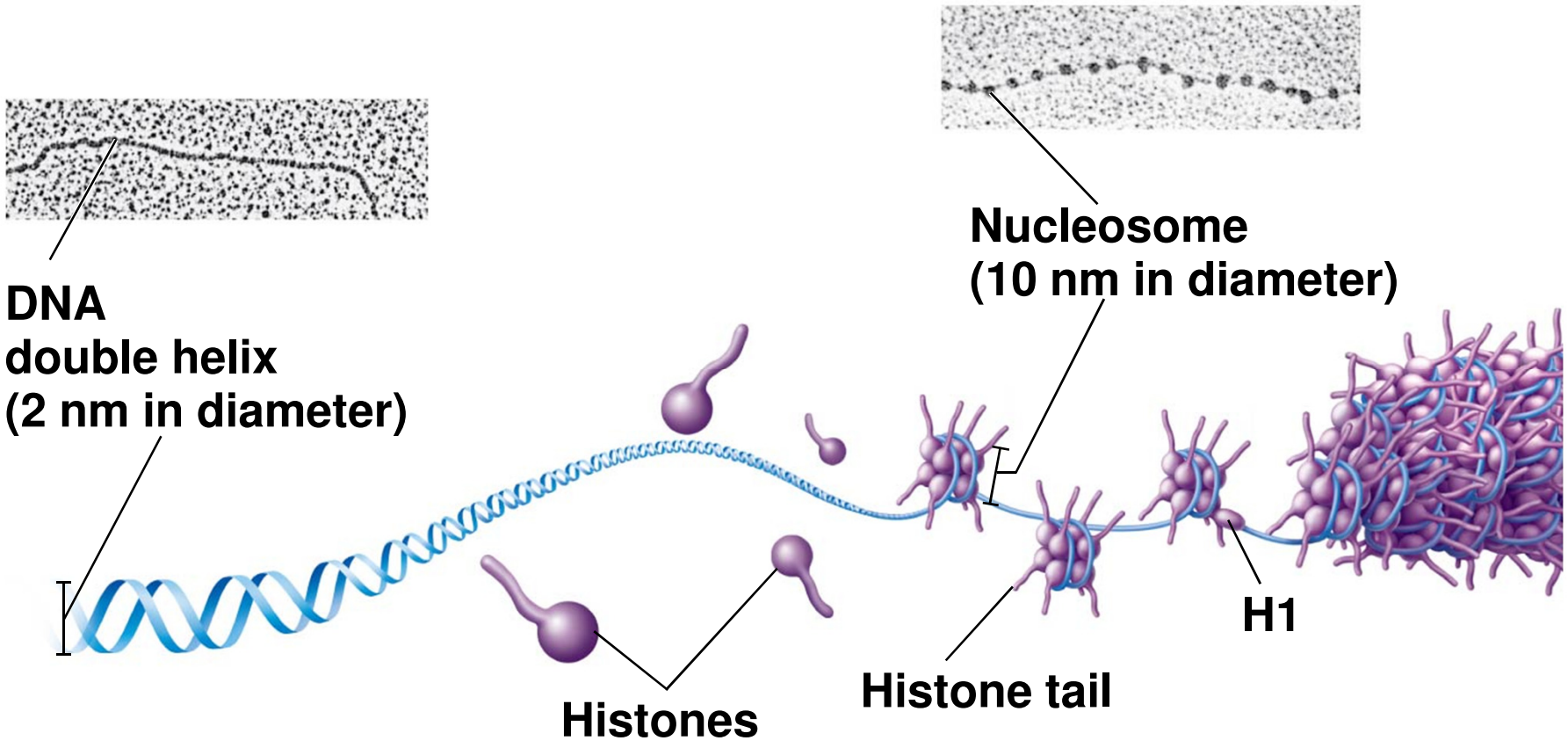
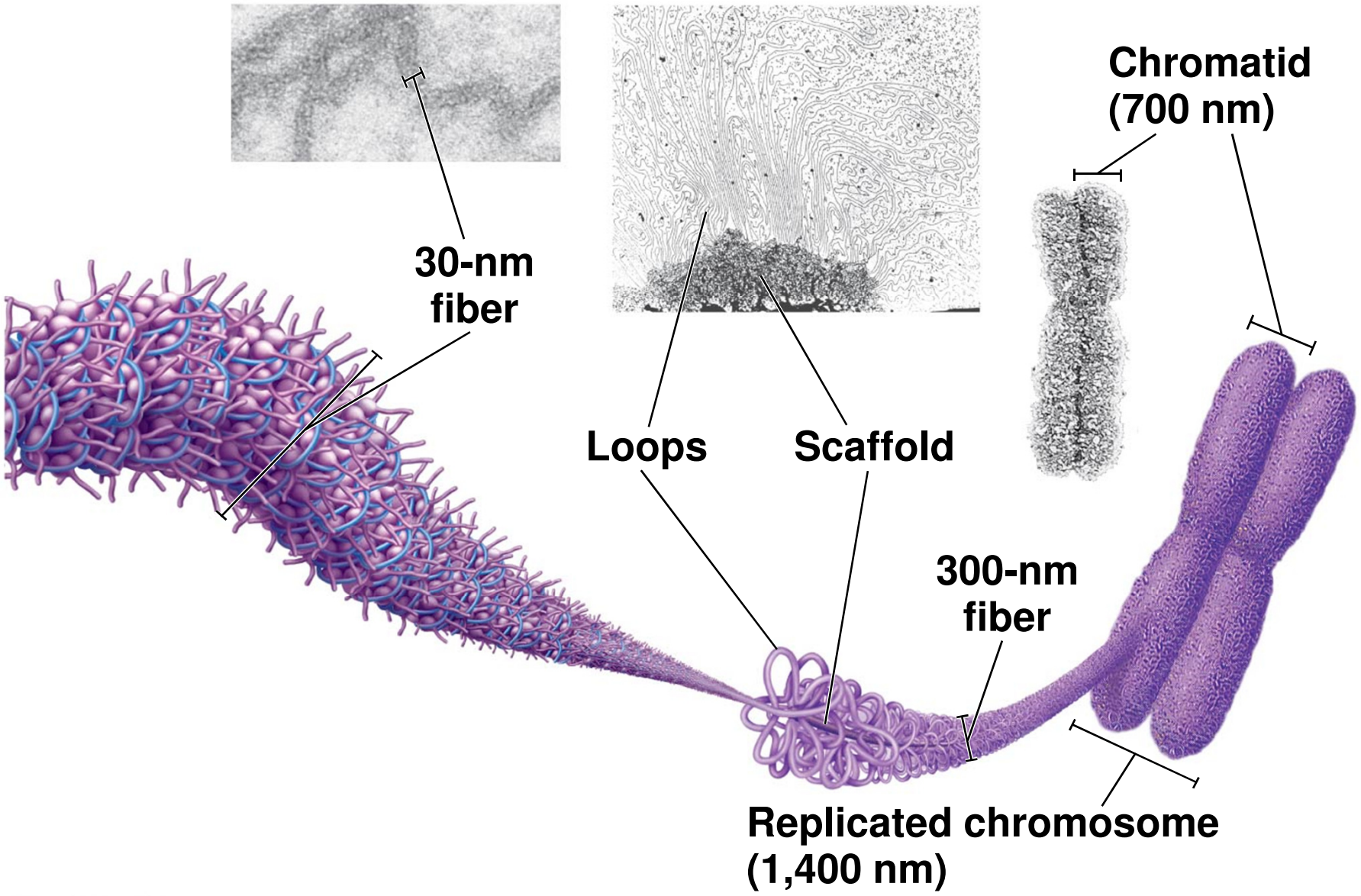
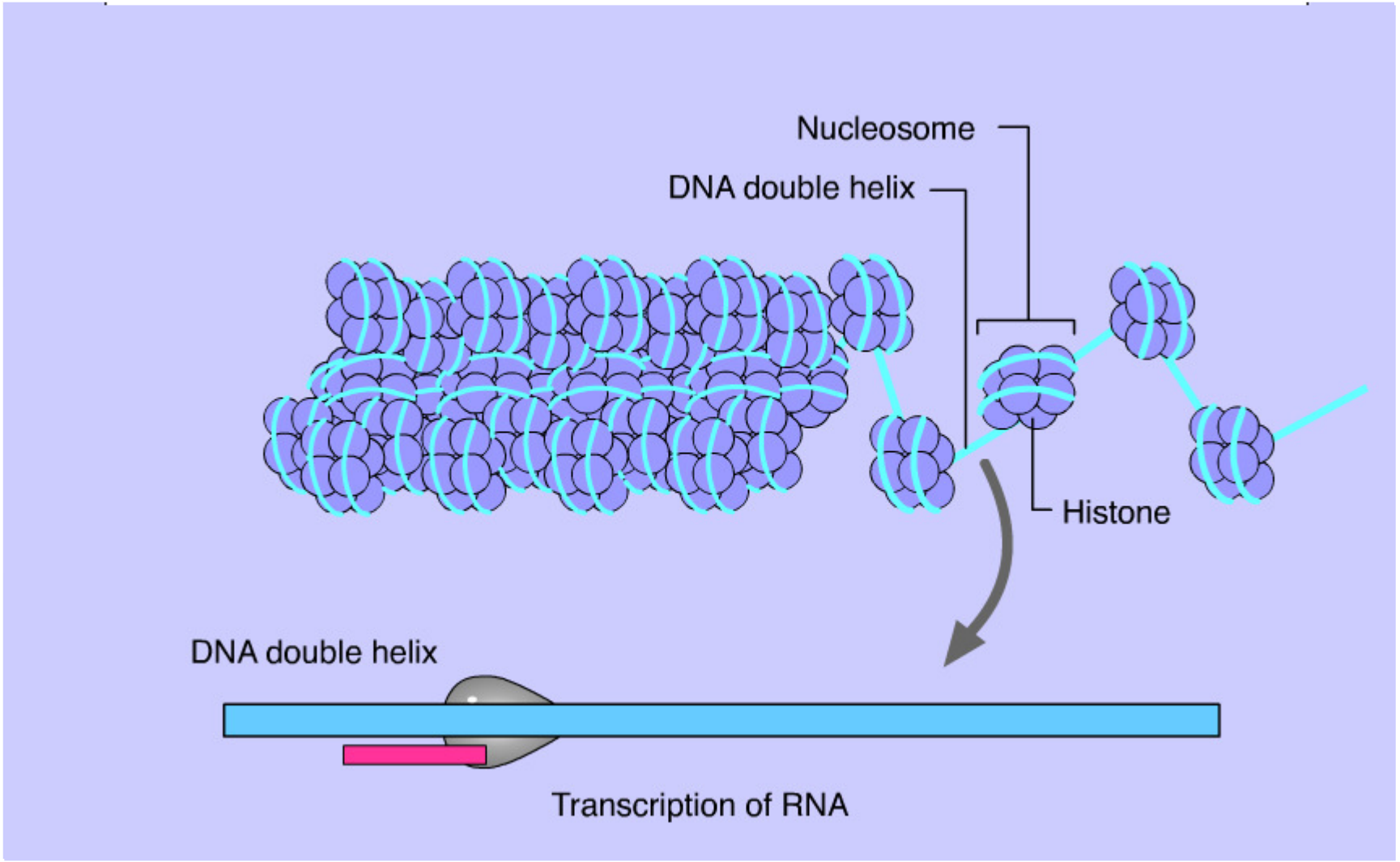


Figure 13.21b



- At interphase, most of the chromatin is compacted into a 30-nm fiber, which is folded further in some areas by looping
- Even during interphase, centromeres and some other parts of chromosomes are highly condensed, similar to metaphase chromosomes
  - This condensed chromatin is called **heterochromatin**
  - The more dispersed, less compacted chromatin is called **euchromatin**

- Dense packing of the heterochromatin makes it largely inaccessible to the machinery responsible for transcribing genetic information
- Looser packed euchromatin makes its DNA accessible to this machinery
  - So the genes can be transcribed
- Chromosomes are dynamic in structure
  - A condensed region may be loosened or modified as needed for various cell processes
  - For example, histones can undergo chemical modifications that result in changes in chromatin organization



Animation: DNA Packing
Right click slide / Select play

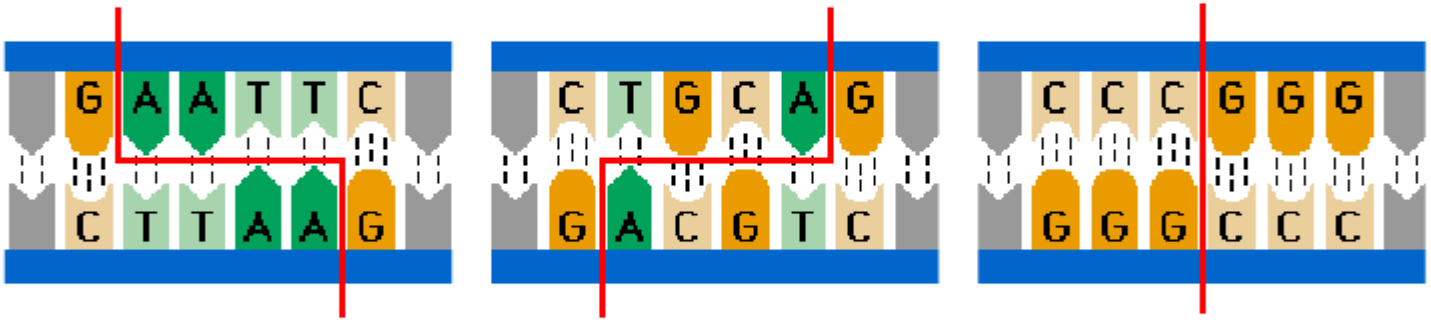
Concept 13.4: Understanding DNA structure and replication makes genetic engineering possible

- Complementary base pairing of DNA is the basis for **nucleic acid hybridization**
  - The base pairing of one strand of a nucleic acid to another, complementary sequence
- Nucleic acid hybridization forms the foundation of virtually every technique used in **genetic engineering**
  - The direct manipulation of genes for practical purposes
  - Applications include agriculture, criminal law, medical research

DNA Cloning: Making Multiple Copies of a Gene or Other DNA Segment

- Genes occupy only a small proportion of chromosomal DNA
  - The rest is noncoding nucleotide sequences
- To work directly with specific genes, scientists prepare well-defined segments of DNA in identical copies
  - Called *DNA cloning*
- Most methods for cloning pieces of DNA in the laboratory share general features

- Many bacteria contain **plasmids**
  - Small circular DNA molecules that replicate separately from the bacterial chromosome
  - Have only a small number of genes
    - May be useful when bacterium is in particular environment but may not be required for survival or reproduction under most conditions
- To clone pieces of DNA, researchers first obtain a plasmid and insert DNA from another source ("foreign DNA") into it
  - The resulting plasmid is called **recombinant DNA**



Animation: Restriction Enzymes
Right click slide / Select play

Figure 13.22a

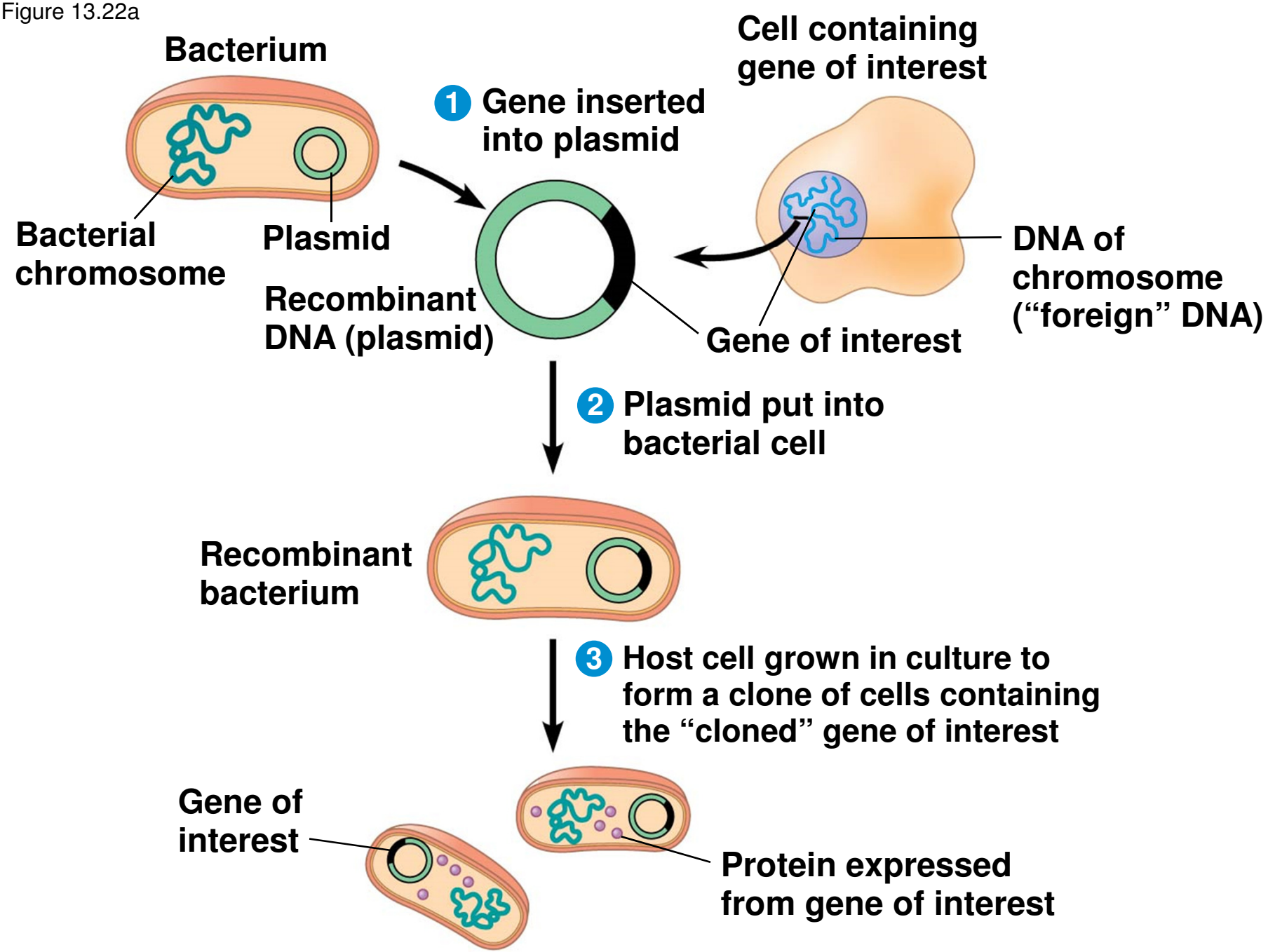
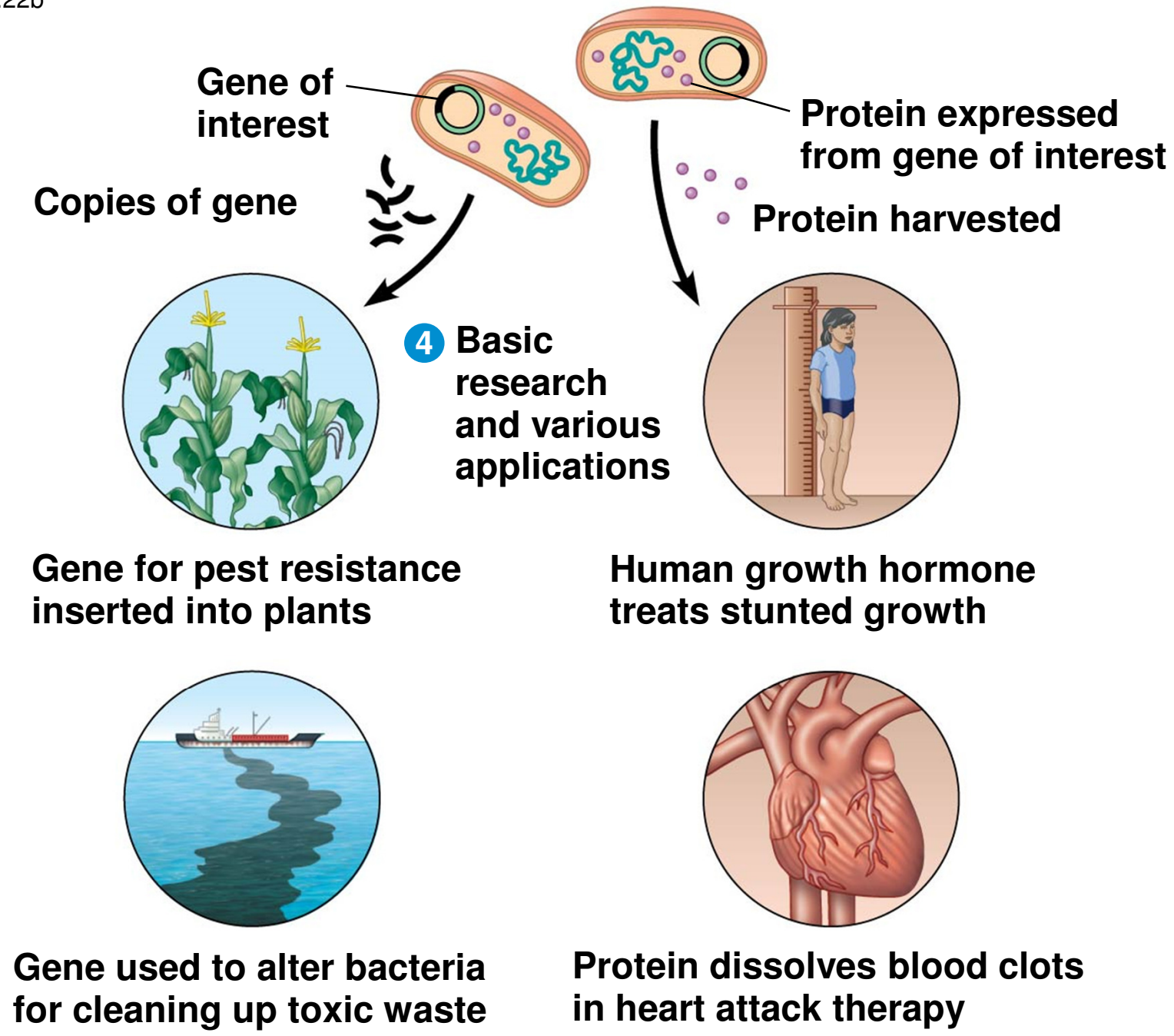


Figure 13.22b

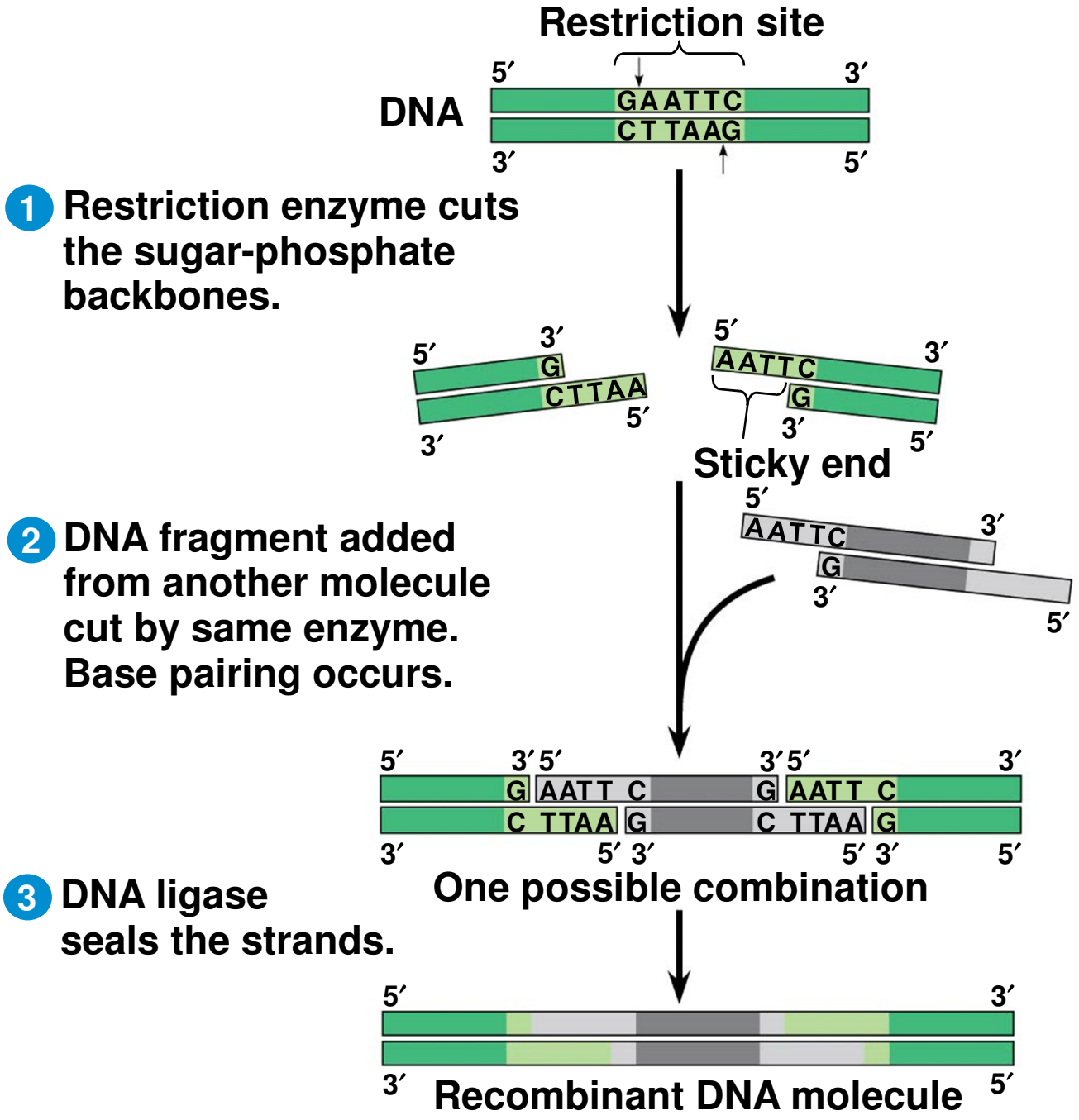


- The production of multiple copies of a single gene is called **gene cloning**
  - Gene cloning is useful to make many copies of (amplify) a gene and to produce a protein product
  - The ability to amplify many copies of a gene is crucial for applications involving a single gene
    - Basic research
    - Endow an organism with a new metabolic trait, such as pest resistance

Using Restriction Enzymes to Make Recombinant DNA

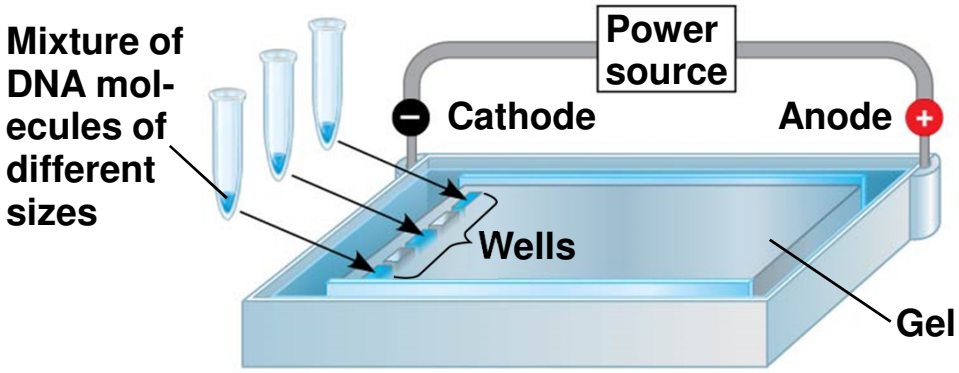
- Bacterial **restriction enzymes** cut DNA molecules at specific DNA sequences called **restriction sites**
  - Each restriction enzyme if very SPECIFIC
  - Cut both DNA strands at precise points within the restriction site
  - Most restriction sites are symmetric
    - The sequence of nucleotides is the same on both strands when read in the 5’ to 3’ direction

Figure 13.23-3

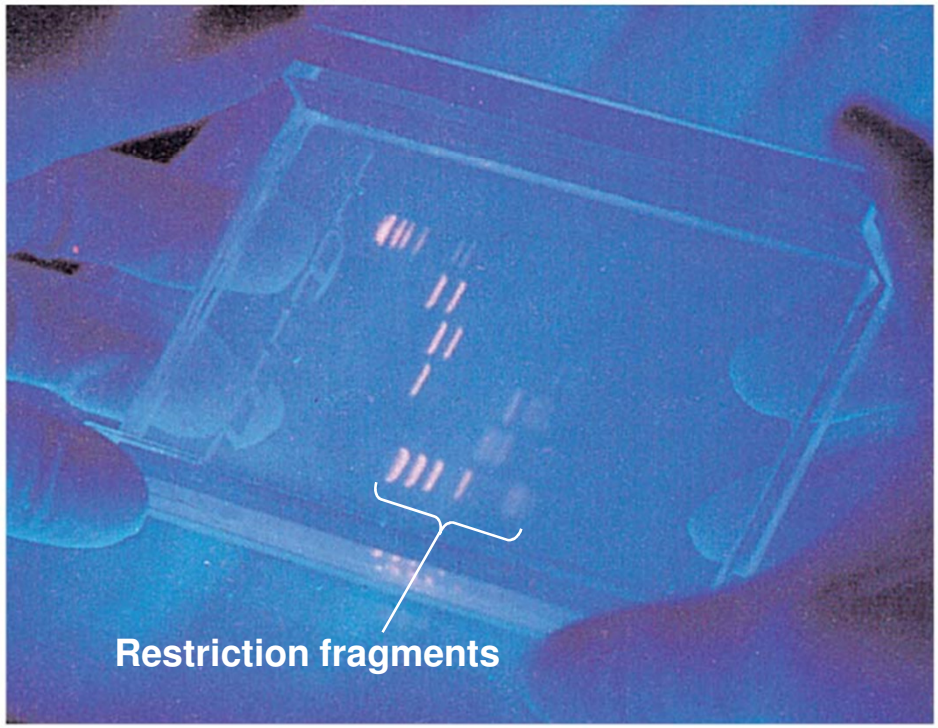


- A restriction enzyme usually makes many cuts, yielding **restriction fragments**
  - All copies of a particular DNA molecule always yield the same set of restriction fragments when exposed to the same restriction enzyme
- To see the fragments produced by cutting DNA molecules with restriction enzymes, researchers use **gel electrophoresis**
  - This technique separates a mixture of nucleic acid fragments based on length

Figure 13.24



(a) Negatively charged DNA molecules will move toward the positive electrode.



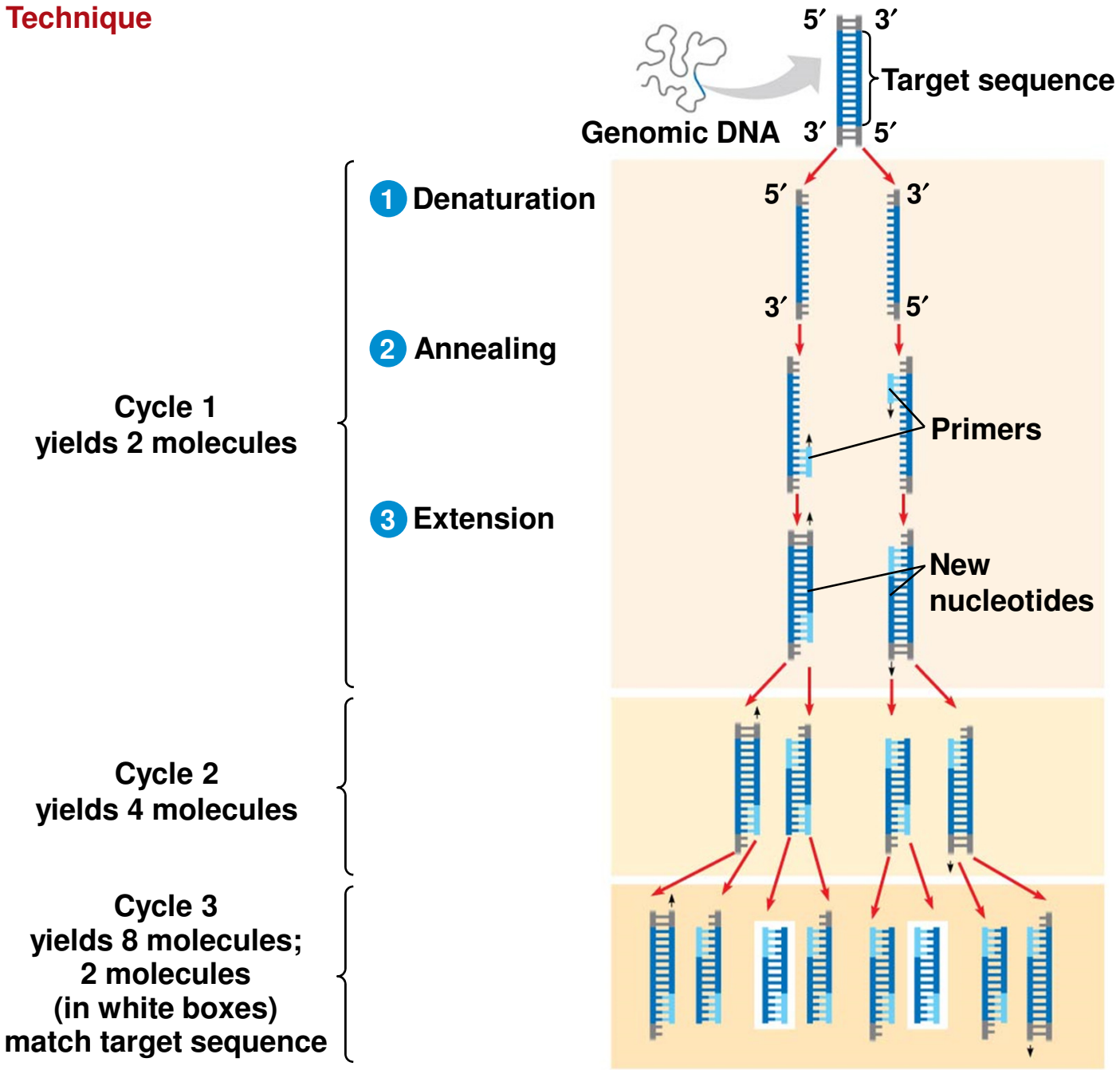
(b) Shorter molecules are impeded less than longer ones, so they move faster through the gel.

- The most useful restriction enzymes cleave the DNA in a staggered manner to produce **sticky ends**
  - Sticky ends can bond with complementary sticky ends of other fragments
  - They are temporary associations but can be made permanent by DNA ligase
    - Can close the sugar-phosphate backbones of DNA strands
- In gene cloning, the original plasmid is called a **cloning vector**
  - A DNA molecule that can carry foreign DNA into a host cell and replicate there

Amplifying DNA *in Vitro*: The Polymerase Chain Reaction (PCR) and Its Use in Cloning

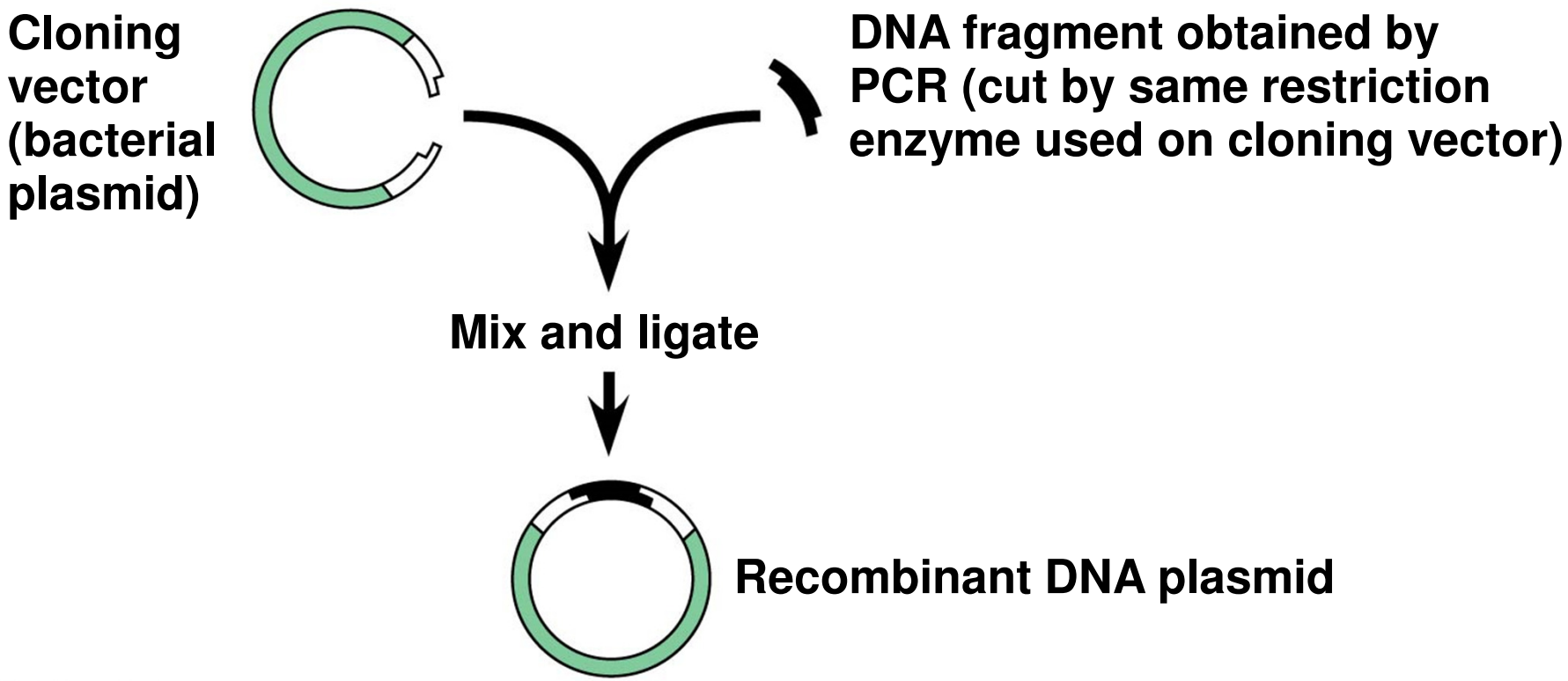
- The **polymerase chain reaction**, **PCR**, can produce many copies of a specific target segment of DNA
- A three-step cycle brings about a chain reaction that produces an exponentially growing population of identical DNA molecules
  1. Denaturation: Heated to separate the DNA strands
  2. Annealing: Cooled to allow primers to form hydrogen bonds with ends of target sequence
  3. Extension: DNA polymerase adds nucleotides to 3' end of each primer
- The key to PCR is an unusual, heat-stable DNA polymerase called Taq polymerase

Figure 13.25


Technique
Genomic DNA
5′
3′
Target sequence
3′
5′
1 Denaturation
2 Annealing
3 Extension
Primers
New nucleotides
Cycle 1 yields 2 molecules
Cycle 2 yields 4 molecules
Cycle 3 yields 8 molecules; 2 molecules (in white boxes) match target sequence

- PCR is speedy and very specific
  - But PCR amplification alone cannot substitute for gene cloning in cells
  - Instead, PCR is used to provide the specific DNA fragment to be cloned
- PCR primers are synthesized to include a restriction site that matches the site in the cloning vector
- The fragment and vector are cut and ligated together

Figure 13.26


Cloning vector (bacterial plasmid)
DNA fragment obtained by PCR (cut by same restriction enzyme used on cloning vector)
Mix and ligate
Recombinant DNA plasmid

- Impact of PCR
  - Been used to amplify DNA from a wide variety of sources
    - 40,000 year old frozen wooly mammoth
    - Fingerprints or tiny amounts of blood/tissue/etc. found at crime scenes
    - Single embryonic cells for rapid prediagnosis of genetic disorders
    - Cells infected with viruses that are difficult to detect

DNA Sequencing

- Once a gene is cloned, complementary base pairing can be exploited to determine the gene's complete nucleotide sequence
  - This process is called **DNA sequencing**
- "Next-generation" sequencing techniques, developed in the last ten years, are rapid and inexpensive
  - They sequence by synthesizing the complementary strand of a single, immobilized template strand